5-2019

Using Self-regulation to Predict Preschoolers' Symptomology of Disruptive Behavior Disorders

Andrew White
University of Nebraska - Lincoln, awhite8@huskers.unl.edu

USING SELF-REGULATION TO PREDICT PRESCHOOLERS' SYMPTOMOLOGY OF DISRUPTIVE BEHAVIOR DISORDERS

by

Andrew S. White

A DISSERTATION

Presented to the Faculty of

The Graduate College of the University of Nebraska

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

Major: Psychological Studies in Education

(School Psychology)

Under the Supervision of Professor Susan M. Sheridan

Lincoln, Nebraska

May, 2019

USING SELF-REGULATION TO PREDICT PRESCHOOLERS' SYMPTOMOLOGY OF DISRUPTIVE BEHAVIOR DISORDERS

Andrew S. White, Ph.D.

University of Nebraska, 2019

Advisor: Susan M. Sheridan

The use of brief rating forms completed by caregivers to identify children at-risk for developing behavioral disorders is common (Lane et al., 2009). However, identifying a behavioral measure assessing child-level variables (i.e., temperamental traits) which predict later behavioral concerns has potential to improve universal screening practices in the context of a multi-tiered systems of support (MTSS) framework. Self-regulation (Rothbart & Bates, 2006) is a trait that is related to externalizing problem behaviors (e.g., Espy et al., 2011), and may be useful as a means to predict young children at risk for developing behavioral disorders.

The purpose of this study is to explore the predictive validity of an established measure of self-regulation (the Head-Toes-Knees-Shoulders Task; *HTKS*; McClelland & Cameron, 2012), for clinically elevated externalizing behaviors (identified using clinical rating forms of externalizing behavior). It was hypothesized that assessing a stable, individual trait such as self-regulation could allow for even earlier identification and intervention among at-risk children than may be available with present screening methods. Participants were 24 preschool students and their classroom teachers. The students were administered the *HTKS* in

their schools and their teachers each completed a rating form assessing behavioral problems across three measures (i.e., the Social Skills Improvement System, Achenbach Caregiver-Teacher Report Form, and Conners Early Childhood Behavior Scale).

Surprisingly, this study did not replicate the relationships between self-regulation and behavioral concerns. Correlations between variables suggested positive relationships between the *HTKS* and two of the behavioral measures (i.e., opposite of the hypothesized direction). Multiple linear regression analyses exploring the relationship between continuous criterion and predictor variables were unable to reject the null hypothesis that *HTKS* does not predict behavioral concerns. Further, logistic regression analyses exploring a dichotomous criterion (i.e., the presence or absence of clinically-elevated behavioral problems) also failed to reject the null hypothesis of the model discriminating behavior problem status no better than chance. Follow-up Receiver Operating Characteristics (ROC) curves and comparison of the area-under-the-ROC-curve (AUC) further suggested *HTKS* was not an effective tool for screening in this context. Finally, the study explores its limitations and proposes additional questions for future research.

Acknowledgments

The process of completing this project and the degree that comes with it has simultaneously been one of the more challenging and rewarding experiences I have been able to endure. I am under no illusions that I accomplished any of this by myself, however, and wish to acknowledge a small sample of the vast support system which helped me through this journey.

Throughout my professional career, I have been fortunate to find myself under the mentorship of several brilliant scientists who have challenged and inspired me to be my very best. I especially would like to thank Dr. Susan Sheridan for her guidance as I have navigated the arduous process of this dissertation from start to completion. Dr. Sheridan, I will always be grateful for your willingness to learn with me as I sought to answer a research question that did not fit neatly under your vast expertise. Thank you for agreeing to "adopt" an additional member into your doctoral seminar five years ago and for pushing me from day one to constantly improve myself. Additionally, I would like to thank Dr. Kathy Rudasill for time spent serving as both unofficial and official co-chair of my committee; your expertise and guidance was invaluable. I wish you nothing but the best in your new position. I would also like to extend thanks to my doctoral committee: Dr. Reece Peterson, Dr. Carrie Clark, and Dr. Rafael De Ayala for their thoughtful and challenging feedback. To my first doctoral adviser, Dr. Merilee McCurdy, thank you for your continued support even after you joined a new program. Dr. Gwen Nugent, and Dr. Gina Kunz, thank you for taking a chance on an undergraduate sophomore with no research experience; the

opportunities you gave me at CYFS so early in my training established the trajectory I now find myself on. Thank you to the remaining school psychology faculty at UNL: Dr. Ed Daly, Dr. Sue Swearer, and Dr. Scott Napolitano, for the quality training and mentorship which has prepared me to begin my career. I also wish to thank Dr. Calvin Garbin and Bruce Venner, your passions for teaching psychology to undergraduates and high school students, respectively, instilled a spark in me before I even knew it was there.

Next, I wish to thank my immediate supports—my incredible cohorts. First, my "first coho" at FSU. I cannot picture a better group of friends and colleagues that could have emerged from such an assortment of people. I especially wish to thank Dr. Paige Blankenship, Dr. Jordy Yarnell, Dr. Megan Crisp, Lily Ware, Shawna Enlow, Katelyn Kuchta, and honorary cohort members John Parker and Josh Ware; you are some of my favorite people and you made Florida feel like home. Somehow I was lucky enough to find myself amongst yet another amazing group of people as I joined my second school psychology cohort. Dr. Sonya Bhatia, thank you for your support and countless drop-in office meetings, and for proving to be the best "conference buddy." Thank you also to Dr. Natalie Hoff and Dr. Allen Garcia for always being available for impromptu "peer supervision" as we navigated the past six years. Thank you to Alicia Krueger, we all miss your spark. I also wish to extend special thanks for the invaluable support from other students: Dr. Heather Schwartz, Dr. Scott Fluke, Dr. Tyler Smith, Dr. Jon Sikorski, Dr. Mindy Chadwell, Dr. Zach Myers, Dr. Ana Damme, thank you.

Most importantly, I wish to thank my family. First, my parents, Steve and Laura White, who never seemed to mind if I stayed up a little later to read and who never allowed me a moment of doubt whether I was loved and supported as I pursued my passions. My brother, Nate White, who has always been an exemplar of good work for the work's sake and who humored me as a freshman excitedly telling him about what I learned in my introduction to psychology course. Thank you to my parents-in-law, Allen and Tammy Buss, especially for being available for last-minute childcare or helping with yardwork so I might carve a few extra hours for writing. A seemingly impossible debt of thanks is owed to my wife and partner, April White. Thank you for an endless stream of support and patience as I disappeared to the office and for always being available when I just needed to vent my frustrations. In my wildest imagination I could not dream up a better partner to spend my life with, and I am grateful for every day we can spend together. Finally, to my son, Henry, I owe a sense of meaning to this project's completion I never could have fathomed when it began. I love you with all that I am and cannot wait to continue to watch you grow (although I still insist you slow down).

Grant Information

This study was supported by a dissertation grant awarded to the author by the Society for the Study of School Psychology

Table of Contents

CHAPTER 1: INTRODUCTION…….. 1
CHAPTER 2: LITERATURE REVIEW……... 3
Temperamental Self-regulation....4
Screening for Behavioral Risk 8
Measuring Self-regulation....11
The Head-Toes-Knees-Shoulders task 14
Summary and Research Questions....16
CHAPTER 3: METHODS…….. 19
Participants and Setting....19
Study Variables and Measures 21
Self-regulation....21
Externalizing problem behaviors 23
Demographic and control variables 25
Recruitment and Data Collection Procedures 27
Data Analyses....30
CHAPTER 4: RESULTS 35
Preliminary Analyses 35
Controlling for group differences and covariates.....36
Regression Analyses 38
Predicting Problem Behaviors Scale of the SSIS....39
Predicting Externalizing Behaviors Scale of the CBCL (C-TRF) 40
Multiple regression. 40
Logistic regression. 40
Predicting Defiant/Aggressive Behavior Scale of the Conners (EC BEH-T) .42
Regression analyses. 42
Logistic regression. 43
Follow-up/Exploratory Analyses 45
ROC curve analyses. 45
Cost-benefit analyses. 47
CHAPTER 5: DISCUSSION 49
Main Findings 49

Unexpected Results and Interpretation 50
Future Directions.....52
Study Limitations 55
Conclusion.....57
ENDNOTES…….. 60
References…….. 61
APPENDIX A: *HEAD TOES KNEES SHOULDERS SCRIPT*…….. 74
APPENDIX B: *ACHENBACH CAREGIVER-TEACHER REPORT* (*C-TRF)* EXTERNALIZING PROBLEM SCALE…….. 82
APPENDIX C: CONNERS EARLY CHILDHOOD BEHAVIOR FORM-TEACHER REPORT (EC BEH-T) DEFIANT/AGGRESSIVE BEHAVIORS TOTAL SCALE…… 83
APPENDIX D: *SOCIAL SKILLS IMPROVEMENT SYSTEM RATING SCALES TEACHER FORM* (*SSIS*) EXTERNALIZING SUBSCALE…… 84
APPENDIX E: PARENT DEMOGRAPHIC QUESTIONNAIRE……85
APPENDIX F: SCRIPT FOR OBTAINING CHILD ASSENT……91
APPENDIX G GROUP MEAN COMPARISONS……92
Group Mean Differences for Covariates and Head Toes Knees Shoulders 92
Group Mean Differences for Covariates and the CBCL – *C-TRF* 93
Group Mean Differences for Covariates and Conners *EC BEH-T*.....94
Group Mean Differences for Covariates and *SSIS* 95

CHAPTER 1: INTRODUCTION

Young children with behavior disorders begin their schooling on a perilous trajectory toward increased rates of school discipline, alternative education placements, and eventual dropout (Snyder, 2001). This trajectory often manifests in preschool, where children with behavior problems begin to fall behind peers across numerous indicators of school readiness (Montes, Lotczyewski, Halterman, & Hightower, 2012). Paramount to altering this trajectory and minimizing long-term consequences is identifying and intervening with at-risk children as early as possible (Lane et al., 2012). The use of brief rating forms completed by caregivers to identify children at-risk for developing behavioral disorders is common (Lane et al., 2009). However, identifying a behavioral measure assessing child-level variables (i.e., temperamental traits) which predict later behavioral concerns has potential to improve universal screening practices in the context of a multi-tiered systems of support (MTSS) framework. Unfortunately, such a screening measure has yet to be validated. Self-regulation (Rothbart & Bates, 2006) is a trait that is highly related to numerous indicators of school readiness, including externalizing problem behaviors (Espy, Sheffield, Wiebe, Clark, & Moehr, 2011), and may be useful as a means to predict young children at risk for developing behavioral disorders.

The purpose of this study is to explore the predictive validity of an established measure of self-regulation (the Head-Toes-Knees-Shoulders Task; *HTKS*; McClelland & Cameron, 2012), for clinically elevated externalizing behaviors (identified using clinical rating forms of externalizing behavior). It was hypothesized that assessing a stable, individual trait such as self-regulation could allow for even earlier

identification and intervention among at-risk children than may be available with present screening methods. A sample of 24 preschool students were administered the *HTKS* in their schools. Each student was recruited as part of a dyad with their classroom teacher who completed a rating form assessing behavioral problems across three measures (i.e., the Social Skills Improvement System, Achenbach Caregiver-Teacher Report Form, and Conners Early Childhood Behavior Scale).

Surprisingly, this study did not replicate the relationships between self-regulation and behavioral concerns. Correlations between variables suggested positive relationships between the *HTKS* and two of the behavioral measures (i.e., opposite of the hypothesized direction). Multiple linear regression analyses exploring the relationship between continuous criterion and predictor variables as continuous were unable to reject the null hypothesis that *HTKS* does not predict behavioral concerns. Further, logistic regression analyses exploring a dichotomous criterion (i.e., the presence or absence of clinically-elevated behavioral problems) also failed to reject the null hypothesis of the model discriminating behavior problem status no better than chance. Follow-up Receiver Operating Characteristics (ROC) curves and comparison of the area-under-the-ROC-curve (AUC) values were also conducted to anecdotally explore the balance between sensitivity and specificity of the task and describe the overall accuracy of *HTKS* as a predictive task. Evidence from these findings also appeared to suggest *HTKS* was not an effective tool for screening in this context. This study was not without limitations, however, and I explore additional questions which future research should address to further elucidate self-regulation's potential as a screener.

CHAPTER 2: LITERATURE REVIEW

The estimated prevalence of behavior and mental health problems in schools exceeds 20 percent, yet less than one percent of students actually receive special education services in the disability category of emotional disturbance (U.S. Department of Education, 2015). In addition, untreated behavioral disorders are related to detrimental long-term effects including school dropout and higher rates of unemployment, substance abuse, and violence (Mental Health America, 2017; Snyder, 2001). Early identification of students at-risk for emotional and behavioral disorders allows the interruption of this trajectory, with the intent to minimize long-term negative consequences (Lane et al., 2012). Externalizing problem behaviors related to emotional and behavioral disorders (e.g., aggression, impulsivity, hyperactivity) also impede the learning of both students with emotional and behavioral disorders and their classmates, largely through classroom distractions (Kalberg, Lane, Driscoll, & Wehby, 2011). Externalizing problem behaviors are especially detrimental for preschoolers' transition to kindergarten (Rimm-Kauffman, Pianta, & Cox, 2000), placing them on a delayed academic trajectory as they fail to develop skills in attending to instruction (Metcalfe, Harvey, & Laws, 2013). Early identification and intervention is critical to prevent long term negative effects of externalizing behaviors on student learning.

Current screening methods (i.e., teacher-completed rating forms) assessing risk for behavioral problems are susceptible to bias. For example, these forms rely upon the judgment of the rater to identify risk. This may be subject to "halo effects," where raters select scores based on their general perception of the child as positive or negative (Merrell, 2000). Therefore, there is a need to investigate the validity of

measures examining individual characteristics assessed at the student level which may be used to predict risk for developing externalizing behavioral disorders. *Self-regulation,* or the ability to purposefully monitor and modulate one's own behavior and reactions to the environment (Rothbart & Bates, 2006), is an individual characteristic that drives socially appropriate behavior and therefore has potential to be used as a construct to screen for externalizing behavioral disorder risk.

Temperamental Self-regulation

Temperament theory describes the biologically-based individual differences in reactivity and regulation that affect a child's emotional or behavioral response patterns in different environments (Rothbart & Bates, 2006). These individual differences are relatively stable across the lifespan and can be observed beginning in infancy (Rothbart, 2011). Through self-regulation, individuals work to control their involuntary or automatic responses to their environments through strategies such as effortful shifting of attention or inhibition of a response (Rothbart & Bates, 2006). Temperamental self-regulation typically is conceptualized as an overarching construct comprising three primary dimensions: *attentional control,* or the selective shifting or focusing of attention; *inhibitory control,* or overriding a dominant response in favor of a non-dominant response; and *activation control,* or initiating a non-preferred response (Eisenberg, Valiente, & Eggum, 2010). In a classroom setting, *attentional control* may present as a student concentrating on a lecture despite noisy distractions. An example of *inhibitory control* could be an exuberant child resisting their impulse to shout an answer rather than raise their hands. In contrast, *activation control* may present as a

shy child raising their hand to answer a question, despite their preferred response to avoid engaging in a class discussion.

Effortful control (i.e., temperamental self-regulation) represents one of the major theoretical approaches to self-regulation, with the neuropsychological approach to executive function providing the other dominant perspective (Blair & Peters Razza, 2007). These approaches were historically confined to their respective literature bases; however, research has recently established considerable overlap across their conceptualizations of self-regulation (Hofmann, Schmeichel, & Beddeley, 2012). Although the current study adopts a temperamental framework, the primary target for exploration is the overarching construct of self-regulation rather than its dimensions. Therefore, some findings from the neuropsychological tradition of self-regulation are included in this review and conceptualized as assessing the same overarching construct.

Young children entering school settings for the first time face expectations (e.g., sitting still, waiting to be called upon) that test their ability to regulate their behavior in ways they may not have previously experienced. Self-regulation serves as one of the critical noncognitive skills that facilitate later academic achievement. Indeed, self-regulatory abilities have been implicated by experienced kindergarten teachers as the most important skill set for school readiness (Lin, Lawrence, & Gorrell, 2003). In addition, relationships between self-regulatory skills and numerous indicators of school readiness exist, including externalizing problem behaviors (Espy et al., 2011).

Specific behavioral problems identified in correlational research as significantly related to low self-regulatory skills include negative student emotionality (*r* range .23–.24; Ferrier, Bassett, & Denham, 2014), off-task behavior adjusted for social competence (*r* = .24; Blair & Peters, 2003), and classroom adjustment (*r* = .19; Denham, Bassett, Sirotkin, Brown, & Morris, 2015). In addition, meta-analytic research comparing self-regulatory skills among children with and without ADHD revealed group differences in performance on self-regulatory tasks, such that children demonstrating ADHD symptomology present significant self-regulatory impairment compared to those without ADHD (*d* range .46–.69; Wilcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Children's reactive aggression (i.e., impulsive aggressive behavior in response to situational stressors) may be even more highly correlated with self-regulation than other identified behavioral problems (*r* = .48; White, Jarrett, & Ollendick, 2013), a finding that emphasizes the importance of self-regulation in children's classroom behaviors and their interactions with peers.

The relationship between performance on self-regulation tasks and problem behaviors has been further explored via longitudinal research. Hughes and Ensor (2008) utilized a longitudinal design to assess the causality of self-regulation skills and problem behaviors from children ages 3 to 4 years. The researchers employed a battery of self-regulation tasks and an aggregate "problem behaviors" score derived from multiple measures of problem behaviors (i.e., rating forms and structured observations of children interacting with parents and in the classroom), and hypothesized that early problem behaviors would constrain self-regulatory development. They found that self-regulation predicted later problem behaviors when they controlled for initial levels of

problem behaviors ($\beta = .46$, $p < .01$), suggesting that early deficits in self-regulation are significant predictors of later problem behaviors.

Espy and colleagues (2011) further elucidated the predictive ability of preschoolers' self-regulation for problem behaviors using advanced modeling techniques. First, the authors identified latent "problem behavior" factor structures via confirmatory factor analysis. Next, these problem behavior factors were fit into structural equation models to be predicted by a composite self-regulation score obtained from a battery of self-regulation performance tasks. This methodology allowed the evaluation of relationship of self-regulation with multiple problem behavior factors, rather than as a single broad construct. The final model included four latent problem behaviors: hyperactivity, attention problems, disinhibition, and emotion dysregulation. The resulting paths between self-regulation and hyperactivity, attention problems, and inhibition were significant in magnitude ($\lambda = -.42$, $-.55$, and $-.48$, respectively; $p < .05$), and the path to emotional dysregulation behaviors was marginal in size but not significant ($\lambda = -.22$). The authors' results provided powerful evidence of the relationship between self-regulation and externalizing problem behaviors, and suggested that laboratory tasks of self-regulation appear to assess the same processes of control that appear in the emergence of disruptive behavior disorders. Therefore, screening for problems in self-regulation early in school may provide a framework for the early identification and intervention of regulatory processes underlying behavioral problems.

Screening for Behavioral Risk

Whereas screening for self-regulatory concerns presents a proposed approach for assessing the mechanisms underlying behavioral problems, extant literature exploring behavioral screeners has instead emphasized raters' perceptions of emergent problem behaviors as they manifest. Common practice in behavioral screening uses brief rating forms completed by teachers or parents to identify youths most at-risk for developing behavior disorders (Lane et al., 2009). Behavioral screeners in the schools are characterized by their ability to identify early symptoms that signify the risk of eventual receipt in special education services, prior to warranting a DSM-V diagnosis (Kamphaus, Reynolds, & Dever, 2014). Among preschoolers this is especially challenging given the fact that many disruptive behaviors (i.e., tantrums, noncompliance) are normative in this age range (Breitenstein, Hill, & Gross, 2009). However, ample evidence suggests that clinically significant disruptive behavior disorders can be meaningfully identified among preschoolers, and waiting to intervene only places children on more perilous trajectories toward increasingly severe conduct concerns later in life (e.g., Loeber & Farrington, 2000; Loeber, Burke, Lahey, Winters, & Zera, 2000).

The growing acceptance and implementation of multi-tiered systems of support in schools (MTSS) has highlighted the importance of identifying valid screening tools for behavioral concerns (Kilgus, Reinke, & Jimerson, 2015). Originally conceptualized within the public health model, MTSS models provide a framework of service delivery emphasizing early intervention informed by ongoing evaluation of student need for and response to services (Doll & Cummings, 2008). MTSS models feature

increasingly intensive services delivered across tiers of support, such that children who are not responding to the interventions present at one tier are advanced to the next tier to receive more intensive services. For example, students who do not respond to a school's universal supports (Tier 1) who possess risk for later problems advance to Tier 2, where they receive secondary supports such as targeted small-group interventions. Students at high risk for additional problems who did not respond to Tiers 1 or 2 may then be moved to the tertiary supports of Tier 3. Tier 3 services are characterized by interventions which are intensive and individualized to the student. In addition to tiers of intervention service delivery, MTSS is also characterized by its use of increasingly intensive assessment of students across tiers. At the Tier 1 level (Universal), all students may be administered highly specific, brief screening measures designed to identify students most likely to require more intensive services. Such procedures are commonplace in the monitoring of academic progress (Lane et al., 2011). Universal screening of academic skills among younger children typically targets emergent skills (e.g., oral reading fluency as a proxy for later reading comprehension) to promote earlier intervention (Goffreda, Diperna, & Pederson, 2009). Teachers generally have less experience screening for behavioral concerns (Lane et al., 2012).

The use of multiple raters (e.g., teachers and parents) for assessing behavior problems may appear to provide a solution to the problem of less reliable teacher reports early in the school year; however, parent and teacher ratings of the same behaviors have historically demonstrated poor correlations with each other (Achenbach, McConaughy, & Howell, 1987; Rudasill et al., 2014), leading to

questions of whether behavioral rating forms were measuring behavioral constructs differently, depending on the rater (McConaughy & Ritter, 1995). Konold, Walthall, and Pianta (2004) explored this further via multigroup confirmatory factor analytic procedures where models were built using parent and teacher ratings from the *Child Behavior Checklist* (Achenbach, 1991a), with invariance compared across models. Konold and colleagues' results suggested that the factor loadings across constructs onto scale scores were indistinguishable across raters despite replicating differences in mean ratings. That is, the measures appear to assess behaviors as designed but may be context-specific, as observed by the raters. Therefore, behavioral screeners in the schools appear most valid when used by those interacting with the children in the school context, where the children are most likely to present the behaviors of concern.

Standard practice appears to present two primary options for identifying children at-risk for behavioral problems. First, schools may wait for students to demonstrate behavioral concerns at sufficiently disruptive levels to necessitate referral (e.g., through tracking indicators of disruptive classroom behaviors such as office disciplinary referrals or suspensions). However, waiting for a sufficient pattern of disruptive behaviors to emerge places students at risk of ingraining themselves with behaviors which likely would have been responsive to early intervention (Gresham, 2007). A need exists, therefore, for screening measures to be deployed in schools which allow educators to identify and intervene upon potentially troublesome behaviors before they can escalate further. Universal screening practices such as those in MTSS provide a second option for identifying children at risk. All students are

assessed using validated measures predictive of their risk for demonstrating later behavior concerns.

Current screening methodology in MTSS frameworks often include teacher-completed behavioral screening measures to identify risk. Missing from this methodology is a validated measure of emergent behaviors which are predictive of significant disruptive behaviors (analogous to the literacy skills assessed as proxies for reading risk). An objective measure of student-level mechanisms underlying externalizing behaviors could theoretically be assessed before any behavioral concerns manifest in the classroom, promoting even earlier identification and intervention. This study proposes screening children's temperamental self-regulation as a substrate for emergent behavioral problems.

Measuring Self-regulation

Despite the evidence of the relationship between self-regulation and externalizing behavior problems, the utility of self-regulatory skills assessment to accurately identify students with behavior disorders from those without has yet to be explored. Indeed, the majority of research exploring the assessment of individual self-regulation has used methodology which is unfavorable for screening purposes. The predominant approach to studying temperamental self-regulation has been the use of rating forms such as the *Children's Behavior Questionnaire* (*CBQ*; Putnam & Rothbart, 2006), which are unwieldy and implausible for use as a universal screening measure given their length and time necessary to complete by third-party raters (i.e., teachers or parents).

A second approach to assessing self-regulation is performance tasks, in which children's regulatory abilities are measured with behavioral paradigms that theoretically serve as behavioral substrates of brain structures underlying the direction of attention and regulation of behaviors (e.g., prefrontal cortex and anterior cingulate cortex; Bell & Deater-Deckard, 2007). Most research in performance-based measures of self-regulation has used batteries (Wiebe, Espy, & Charak, 2008) that comprise several tasks purporting to assess the dimensions of self-regulation (i.e., *attentional control, inhibitory control,* and *activation control*). Each task employs a unique approach to assessing self-regulation (i.e., "paradigms"), which seeks to emphasize a given dimension.

Persistence paradigms (e.g., Goldsmith, Reilly, Lemery, Longley, & Prescott, 1993) require continued attention and perseverance through a monotonous task, such as correctly sorting beads into specified containers (Bead Sorting task; Goldsmith et al., 1993), or maintaining effortful attention to remember rules (e.g., Stroop task paradigms requiring the ignoring of salient information in favor of another feature). These paradigms primarily capture the self-regulatory dimension of attentional control. Activation control, in contrast, may be captured through compliance paradigms (e.g., Toy Cleanup; Kochanska, Coy, & Murray, 2001) wherein the child is given directions to initiate a non-preferred activity (such as cleaning up toys) and their latency to comply is measured. These and similar paradigms are typically included in batteries of self-regulation measures; however, the majority of research assessing performance tasks of self-regulation include tasks utilizing inhibition paradigms.

Inhibition paradigms require the child to inhibit a preferred response in favor of a non-preferred response, and a number of tasks have emerged seeking to capture the construct of inhibitory control. Specifically, popular methods of assessment have included delaying tasks (e.g., Mischel, Shoda, & Rodriguez, 1989), go/no-go tasks (e.g., Carlson & Moses, 2001), and conflict tasks (e.g., Passler, Isaac, & Hynd, 1985). "Delaying" tasks (Mischel et al., 1989) require children to resist a temptation (e.g., eating a marshmallow) in favor of receiving a better reward (e.g., two marshmallows) at a later time (thereby inhibiting their preferred response to engage with the temptation immediately). "Go/no-go" tasks require children to provide a response under certain conditions but withhold responding under others (e.g., the "Simon Says" game where children must follow directions only when they are preceded by the phrase "Simon Says"). These "go/no-go" tasks measure inhibition of the children's preponderance to respond to every prompt, rather than initiating only under the requisite conditions. Finally, "conflict" tasks require children to provide a non-intuitive response over the intuitive response. The Grass/Snow task (Carlson & Moses, 2001; adapted from Passler et al., 1985) is a conflict measure which requires children to point to a white or green piece of paper when the evaluator says "grass" or "snow," respectively, pointing to the paper opposite of the word's color. Conflict tasks assess inhibition by requiring children to respond in a manner which is contrary to their natural inclination (e.g., pointing to the opposite color of the cue word in Grass/Snow).

The composite of these batteries represents the overarching construct of self-regulation. However, any performance measure of self-regulation necessarily assesses all dimensions of self-regulation to varying degrees, and can never purely assess a

single dimension (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Miyake and colleagues (2000) used structural equation modeling to compare models assessing dimensions of self-regulation as both a three-factor model (i.e., self-regulation as a unitary construct) and a "three independent factors" model (i.e., self-regulation as three orthogonal dimensions). The authors found that although the dimensions are separable, they possess an underlying commonality which precludes pure assessment of one dimension independent of the others. Assessing unique self-regulatory dimensions is even more challenging among younger children, to the extent that self-regulation cannot be meaningfully differentiated into separate components and instead is best interpreted only as a single construct (Carlson, 2005). Given this, the assessment of self-regulation via large batteries, which can be time- and resource-prohibitive in a number of settings (McClelland et al., 2014), represents a much less ecologically-sensitive approach (Shaul & Schwartz, 2014) thus creating a need for researchers to identify single tasks which accurately capture the construct of self-regulation.

The Head-Toes-Knees-Shoulders task

An emerging task for measuring self-regulation, the Head-Toes-Knees-Shoulders (*HTKS*; McClelland & Cameron, 2012), has been developed and evaluated as a single task of self-regulation (McClelland et al., 2014). *HTKS* is quick (i.e., less than five minutes) and administered directly to a child. *HTKS* requires the child to respond in an unusual manner to commands from the instructor; for example, touching his or her toes when instructed to "touch your head," his or her knees when instructed to "touch your shoulders," and vice versa. *HTKS* integrates the dimensions of self-

regulation (through the expectations of paying attention to directions, remembering rules, executing a non-dominant response, and adapting to shifts in rules as knees-shoulders are added) into a task conceptualized as assessing the construct of behavioral self-regulation (McClelland et al., 2014). In addition, *HTKS* includes rule changes which increase the task's difficulty, reducing ceiling effects present in many other measures of self-regulation. Early research examining *HTKS* sought to use the task to explore the relationship between the overarching construct of self-regulation and emergent academic skills, extending earlier research which suggested that dimensions of preschoolers' self-regulation (i.e., attentional control) predict later academic outcomes including math and reading achievement (NICHD Early Child Care Research Network, 2003). McClelland and colleagues' (2007) study showed that growth in *HTKS* predicts academic outcomes including small but significant effect sizes for preschoolers' emerging literacy, vocabulary, and math skills, such that preschool children who demonstrated greater improvements in *HTKS* scores from fall to spring also demonstrated greater gains in emerging literacy ($d = .09$, $p < .05$), vocabulary ($d = .15$, $p < .05$), and early math ($d = .09$, $p < .05$).

Another study found significant regression coefficients for *HTKS* for preschoolers' mathematics ($\beta = 0.14$) and kindergartner's mathematics ($\beta = 0.15$), early literacy ($\beta = 0.17$), and vocabulary ($\beta = 0.16$; McClelland et al., 2014), but this relationship is mediated by students' problem behaviors ($r^2 = .21$) and social skills ($r^2 = .30$; Montroy, Bowles, Skibbe, & Foster, 2014). Thus, self-regulation measured by *HTKS* appears to predict problem behaviors or social skills that, in turn, predict academic outcomes. The current study elected to emphasize the relationship between

self-regulation and problem behaviors, given the impact of problem behaviors on students' long-term trajectories and problem behaviors' interference with the classroom experiences of their peers. In addition, the use of *HTKS* to assess self-regulation sought to contribute to the dearth of literature exploring the use of individual performance tasks within a screening context.

Summary and Research Questions

As schools continue to promote universal behavioral screening as part of multi-tiered systems of support, empirical studies must identify effective means for evaluating individuals as early as possible to intervene quickly and optimize children's academic and social trajectory. In addition to predicting severe long-term negative consequences for children with behavior problems, classroom behavioral problems also tend to interfere with the learning of the child's peers.

A measure of temperamental self-regulation, such as *HTKS*, may allow the individual assessment of a mechanism of behavior which precedes externalizing behavior disorders. Identifying such a measure is crucial, given the long-term trajectory for children who remain unidentified for developing behavior disorders (Metcalfe et al., 2013). Typically, educators rely on rating forms completed by adults based on the behaviors they have witnessed from the child being evaluated.

The purpose of this study was to determine the *HTKS* task's predictive validity as a behavioral screener for children at-risk of demonstrating clinically elevated externalizing behavior disorders. This measure allows researchers to identify individual differences across children in a temperamental trait (i.e., self-regulation) that has been closely linked to behavioral difficulties. The long-term goal of this line

of research is to improve the early identification of preschoolers with externalizing behavior disorders and thereby ameliorate the development of significant, detrimental trajectories including alternative education placements, school drop-out, or unemployment (Mental Health America, 2017; Snyder, 2001).

The project is innovative because it proposed a new and simple direct method for behavioral screening, and utilizes new analytic methods to explore the evidence of predictive validity for the *HTKS* task in a novel context (i.e., as a behavioral screener). Such efforts will advance theoretical and empirical understandings of self-regulation's role in externalizing problem behaviors and will directly inform practice. Specifically, self-regulatory skills are significantly, negatively related to externalizing problem behaviors (e.g., Espy et al., 2011). However, researchers exploring this relationship have relied upon rating forms or batteries of tasks assessing self-regulation administered in laboratory settings. In contrast, this study assessed self-regulation using a single performance measure administered in the schools, better approximating universal screening conditions. The research sought to establish the groundwork for a line of research and practice incorporating assessment of self-regulation, a construct critical to school success, in early behavioral screening contexts. This would theoretically allow for earlier identification and intervention among at-risk children than may be available with present screening methods. Moreover, results of this study will inform school-based practitioners as they seek to incorporate efficient universal screeners within multi-tiered systems of support.

This study sought to answer two research questions. The first question the study asked was *"Does temperamental self-regulation predict externalizing behaviors?"*

It was hypothesized that self-regulation would negatively predict externalizing behaviors, such that each one-unit increase in self-regulation would predict a corresponding decrease in externalizing behaviors, holding other variables constant.

The second question proposed in the study was *"How well does a performance measure of temperamental self-regulation differentially predict preschoolers who demonstrate clinically elevated externalizing behaviors from those preschoolers who do not?"*

It was hypothesized that children's self-regulatory abilities (as indicated by total scores on the *HTKS*) would accurately identify children at-risk for developing an externalizing behavior disorder (i.e., children with clinically elevated scores on standardized rating measures of behavior disorders), such that children with low self-regulatory scores would demonstrate clinically elevated behavior disorders, and children with high scores on the *HTKS* would not.

CHAPTER 3: METHODS

Participants and Setting

Participants were 24 preschool students (16 boys and 7 girls), drawn from various center-based preschools across the Lincoln and Omaha metropolitan areas and surrounding communities. Students were recruited from preschools in neighborhoods with various demographic profiles, focusing recruitment on parochial and for-profit centers. To control for classroom effects that could emerge from one teacher rating more than one child and potentially violate assumptions of normal error distribution, only one child per teacher was recruited. Exclusionary criteria included children with developmental disabilities, English Language Learners, and children younger than four years or older than five years, eleven months to allow conversion into normative scores on all of the rating forms.

To determine the number of participants required to detect a significant effect, pilot analyses were conducted using an existing dataset from a prior study with permission from that study's primary investigator (Dr. Caron Clark). Although that study did not use identical measures, the researchers collected data on self-regulation using a lab-based, computer administered response inhibition task (*Fish-Shark* task; Wiebe et al., 2012) and the externalizing problems scale score of a teacher-completed rating of behavior (*Achenbach Caregiver-Teacher Report*; *C-TRF;* Achenbach & Rescorla, 2000). The *Fish–Shark* task requires children to click a button to "catch" fish on the screen but must withhold pressing the button when a shark appears on the screen. Children's *C-TRF* externalizing behavior *T*-scores were sorted into two groups, $T \geq 60$ and $T < 60$, indicating clinically elevated externalizing problem

behaviors. Logistic regression analyses using Fish–Shark score to predict the clinically elevated behavior problems status revealed a significant odds ratio (e^{b1} = .115, p = .004). Inputting this odds ratio into G*Power (α = 0.05) revealed output parameters placing the critical z = -1.64 and a required n = 24. The study achieved a final sample size of 24 children. Demographic characteristics of the sample follow (see Table 1).

Table 1

Demographic Characteristics of Sample

	N	*Sample %*	*Mean (SD)*
Age (Months)	23*		57.6 (5.9)
Gender			
Male	16	69.6%	
Female	7	30.4%	
Race/Ethnicity++			
Hispanic/Latino	1	4.2%	
American Indian	1	4.2%	
Asian	1	4.2%	
Black	2	8.3%	
Hawaiian/Pacific Islander	1	4.2%	
White	21	87.5%	
Previous Center-based Preschool Experience			
None	4	16.7%	
One semester	5	20.8%	
Two semesters	2	8.3%	
Four semesters	9	37.5%	
Five or more semesters	4	16.67%	
Socioeconomic Status			
Parent Combined Income			
$45,001-50,000	1	4.2%	
$70,001-$75,000	1	4.2%	
$75,001-$100,000	5	20.8%	
$100,001-$200,000	15	62.5%	
$200,001 or more	1	4.2%	
Prefer not to answer	1	4.2%	
Maternal Highest Education			
Some College	1	4.2%	
Associate's Degree	2	8.3%	
4-year College Degree	9	37.5%	
Some Graduate School	5	20.8%	
Graduate/Professional Degree	7	29.2%	

Note. *One parent did not report their child's birthdate. ++Percentages total greater than 100% because families were allowed to identify more than one race

Study Variables and Measures

Multiple measures were used to assess each child's demonstration of externalizing behaviors and self-regulation ability. Multiple rating forms were included, rather than a single measure of externalizing behavior problems, in order to strengthen evidence of the validity of the self-regulation task's ability to predict behavior problems across scales with similar but not perfectly overlapping constructs of behavior disorders. Additionally, relevant covariates (e.g., children's demographic characteristics) were considered when examining the associations between children's self-regulation and externalizing behavior disorders.

Self-regulation. The predictor variable in this study is preschoolers' self-regulation. Self-regulation requires demonstration of *attentional control* (selective shifting or focusing of attention), *inhibitory control* (overriding a dominant response in favor of a non-dominant response), and *activation control* (initiating a non-preferred response) (Eisenberg, Valiente, & Eggum, 2010). The behavioral self-regulation task individually administered to child participants was the Head-Toes-Knees-Shoulders task (*HTKS*; McClelland & Cameron, 2012; see Appendix A), a structured observation of self-regulation. *HTKS* is a game-like activity that requires the child to respond in an unusual manner to commands from the instructor (e.g., touching his or her toes when instructed to touch their head, or their knees when instructed to touch their shoulders). The child is first assessed for understanding the names for their head, toes, knees, and shoulders; if they point correctly to each body part, they are instructed to "be a little silly and do the opposite of what I say" and the assessment begins. The *HTKS* task comprises three parts that receive scores: Practice Items, Part I, and Part II. The

Practice Items are four items that require the child to practice pointing to the opposite body part than instructed (head or toes), receiving corrective prompts for incorrect responses. If the child answers three of the four practice items incorrectly, testing is discontinued. If the child answers at least two Practice Items correctly, the child then advances to Part I. In Part I, the child is asked to continue playing the game and doing the opposite of what is said for ten trials without receiving corrective prompts. If the child correctly responds to five or more items in Part I of the task (i.e., points to head when instructed to point to toes and vice versa), Part II is initiated. In Part II, the child is given ten more items including both head/toes and knees/shoulders commands. Each movement in the Practice Items, Part I, and Part II is scored on a three-point scale ("0" for failing to touch the opposite body part, "1" for a self-correction without prompting, "2" for immediately touching the opposite body part). The final score is the sum of the scores across the practice items and Parts I and II (40 possible points). Raw scores are interpreted, and although norms have not been developed, previous studies have reported mean scores of 17.38 for 4.5 year-olds (McClelland et al., 2014), 24.73 for 5 year-olds (McClelland et al., 2014), and 26.8 and 27.5 (Ponitz, McClelland, Matthews, & Morrison, 2009; Matthews, Ponitz, & Morrison, 2009) for 5.5 year-old preschoolers. *HTKS* requires approximately five minutes to administer to most children, and has established high interrater reliability with preschool samples (κ = .90; McClelland et al., 2007), as well as strong construct validity based on parent and teacher reports of self-regulatory ability (Ponitz et al., 2009) and other performance tasks of self-regulation including flexibility (*r* range 0.46–0.56), working memory (*r*

range 0.38–0.41), response inhibition (*r* range 0.13–0.40), and go/no-go (*r* range 0.38–0.54) paradigms (McClelland et al., 2014).

Externalizing problem behaviors. The *criterion variable* in this study was clinically elevated externalizing behavior problems comprising noncompliance, aggression, and impulsive/hyperactive behaviors (McMahon, 1994). Selected scales from multiple clinical rating forms of problem behaviors were completed by participating teachers to provide more robust evidence of the *HTKS* task's ability to predict problem behaviors. Each form's raw scores were converted to *T*-scores to allow normative interpretation of child problem behavior. To address the research question of predictive utility, child *T*-scores were coded into one of two dichotomous codes: $T \geq 60$ and $T < 60$. A *T*-score of 60 represents the 84th percentile and is typically used as a cutoff for clinically elevated scores on school-based screenings for behavior disorders (DiStefano & Morgan, 2011). Indeed, using *T*-score $\geq$ 60 as the cutoff is a better predictor of a child's likelihood of eventual referral than using higher, "clinically severe" scores (i.e., 70; Achenbach, 1991a, 1991b). Specifically, teachers completed the Externalizing Problem Scale (Attention Problems and Aggressive Behavior) of the *Achenbach Caregiver-Teacher Report* (*C-TRF*, Achenbach & Rescorla, 2000; see Appendix B), the Defiant/Aggressive Behaviors Total Scale of the *Conners Early Childhood Behavior Form-Teacher Report* (*EC BEH-T;* Conners, 2009; see Appendix C); and the Externalizing subscale of the Problem Behaviors Scale of the *Social Skills Improvement System Rating Scales Teacher Form* (*SSIS;* Gresham & Elliott, 2008; see Appendix D). These subscales were selected due to their similar but not identical constructs representing externalizing problem behaviors.

C-TRF. The Child Behavior Checklist, Teacher Report *C-TRF* (Achenbach & Rescorla, 2000) is a 99-item, norm-referenced behavioral rating form for children aged 18 months through 5 years. Internal consistency on this measure averages .80 (Achenbach & Rescorla, 2000). Teachers rate the extent to which behaviors describe the target child using a 3-point (0–2) Likert-type scale ("not true, somewhat or sometimes true, very true or often true"). Participating teachers completed the *Externalizing Problems Scale* of the *C-TRF* (see Appendix B), which comprises 32 items representing symptoms of inattention and aggression, and takes approximately 4 to 5 minutes to complete. Although the *C-TRF* manual does not report internal consistency statistics for the measure, subsequent studies have shown the *Externalizing Problems Scale* of the *C-TRF* to demonstrate acceptable internal consistency ($\alpha = 0.90$; Kristensen, Henriksen, & Bilenberg, 2010). Sample behaviors assessed by the *Externalizing Problem Scale* include not sitting still, defiant behavior, hitting others, and screaming. Items on the *C-TRF* were observed to demonstrate excellent internal consistency in this study ($\alpha = 0.93$).

EC BEH-T. The Conners Early Childhood Behavior Form-Teacher Report *EC BEH-T* (Conners, 2009) is a norm-referenced rating form following a Likert-type scale with ratings from 0–3 ("never, occasionally, often, very frequently"). The *EC BEH-T* is valid for use with children ages 2 to 6 years old. Internal consistency for the *EC BEH-T* is adequate, with $\alpha = .75–.96$ across subscales (Conners, 2009). Construct validity was supported via strong correlations ($r = .66–.93$ across measures) with similar measures of behavior, including the *C-TRF* (Conners, 2009). Participating teachers completed the *Defiant/Aggressive Behaviors Scale* of the *EC-BEH-T* (see

Appendix C), which comprises 18 items and takes approximately 3 to 5 minutes to complete. *The Defiant/Aggressive Behaviors Scale* of the *EC-BEH-T* demonstrates acceptable internal consistency (α = 0.94; Conners, 2009). Sample behaviors assessed by the *Defiant/Aggressive Behaviors Scale* include not following directions, losing temper, arguing, fighting, and bullying. Items on the *EC-BEH-T* were observed to demonstrate good internal consistency in this study (α = 0.89).

SSIS. The final form completed by teachers was the *Problem Behaviors Scale* of the *SSIS* (Gresham & Elliott, 2008), a norm-referenced rating form where raters rate the frequency a given behavior has occurred during the past two months using a 4-point rating scale ("never, seldom, often, or almost always"). The *Problem Behaviors Scale* of the *SSIS* comprises 46 items and can be administered to rate children from 3 to 18 years old. The *SSIS* on the teacher rating form presents adequate internal consistency (α = .75–.96 across age groups; Gresham & Elliott, 2008) and construct validity across similar measures of behavior (e.g., BASC-2; Reynolds & Kamphaus, 2004). Teachers completed the *Externalizing Problem Behaviors Subscale* of the *SSIS* (see Appendix D), which comprises 12 items and takes approximately 2–3 minutes to complete. The *Externalizing Problem Behaviors Subscale* demonstrates acceptable internal consistency for children ages 3–5 years (α = 0.93; Gresham & Elliott, 2008). In this study the items on the *SSIS* demonstrated excellent internal consistency (α = 0.91). Altogether, teachers required approximately 10–20 minutes to complete the provided subscales for their student.

Demographic and control variables. The information packet provided to parents included a brief demographic questionnaire to compare the recruited child

sample to Nebraska preschool demographics (see Appendix E). Specific demographic information collected included child birthdate, ethnicity, race, previous preschool experience, combined parent income, and mother's highest level of education. These demographic data served as covariates in the linear regression model. No effects were found for most covariates and they were excluded during logistic regression analyses to allow the model to remain full rank. Previous preschool experience did have a relationship with the - and was included in analyses for that measure. Covariates were proposed based on research implicating each in the development of children's self-regulation. Age was measured in months to capture the rapid increase in performance on self-regulation tasks which occurs from ages three to five years (Rothbart & Bates, 2006).

Preschool experience was specified to capture the development in self-regulation which appears to be promoted by the structure and demands of the classroom setting (Bronson, 2000). "Previous preschool experience" was defined as students spending the majority of their daytime hours at licensed pre-kindergarten programs (i.e., public school preschool, Head Start program, early education center, and parochial child care centers) as identified by parents on the demographic questionnaire (Appendix E). Parents identified where their children spent the majority of their weeks for four time periods (previous semester, previous summer, previous school year, and any time before then). The resulting scores produced a "previous preschool experience scale" (i.e., 0 = no previous experience to 4 = experience in all four time periods).

Socioeconomic status was included to control for established environmental factors which may promote disparities in prefrontal-dependent cognitive abilities between children from high- and low-income households (Hackman & Farah, 2009). Parents selected their household combined income from provided ranges (see Appendix E). Only one family reported a combined income below $70,000 (Table 1). Maternal education was also identified as a covariate for its role as a proxy for socioeconomic status. Overall this sample did not present with sufficient variability in socioeconomic status; as a result neither income nor maternal education were included as covariates in analyses.

Finally, because differential performance on self-regulation tasks has been connected with children's ethnic minority status (Caughy, Mills, Owen, & Hurst, 2013), ethnicity/race minority status was also proposed as a covariate in the model. Ethnicity was assessed through one question on the demographic form asking parents to identify the child's ethnicity (i.e., "Hispanic" or "Non-Hispanic"). Race was assessed through a multiple response option asking parents to identify the child's race (i.e., "White," "Black or African American," "American Indian/Alaska Native," "Hispanic or Latino," "Asian" "Native Hawaiian or Pacific Islander," "Two or More Races" or "Some other race"). Ethnicity/race minority status was utilized as a dichotomous variable, defined as children identified as "White, non-Hispanic" (non-minority) and all other categories (minority).

Recruitment and Data Collection Procedures

The present study collected data in participating preschool classrooms recruited in the Lincoln and Omaha metropolitan areas. Agency administrators (i.e., principals

or preschool directors) were invited to participate in the study via email or "cold calls" on the telephone. Informed consent to recruit student-teacher dyads was verbally obtained from agency administrators when they were invited to participate in the study. Once administrators approved recruitment to begin, individual teachers were contacted to provide information about the study and receive an invitation to participate. Some sites with multiple classrooms preferred for coordination to occur through the administrator. In these cases data collectors met with the administrator in-person to review procedures and obtained written consent from teachers on the day of assessment. Teachers were provided additional information with details of the project, including the phone number and email address of the researcher and committee co-chairs for follow-up communication and questions. Teachers' written consent to participate was obtained, and they were provided with brief packets to send home with their students containing information about the study and its benefits and procedures for parents, author contact information, and a written consent for parents to sign and return to the school. Packets were sent home with all eligible students; that is, students ages 4:00-5:11 years, who had not been identified with developmental delays, and whose primary language was English. Families were allowed between one to two weeks to return the packets to the preschool, and classes with at least one child receiving consent were scheduled to participate. Among classrooms with multiple returned consents, one child was randomly selected the day of assessment using a random number generator in Excel based to sort their participant ID numbers (i.e., cell *= rand( )*). Packets were organized in ascending order based on the randomly

generated numbers, and the lowest-numbered child with completed parental consent and demographic packets was selected to participate.

Data collectors included the primary investigator and two trained undergraduate students hired for the study. Training included didactic and live administrations of the *HTKS* task until 90% reliability was obtained. Data collectors maintained reliability greater than 90% throughout assessments across five co-coded assessments (average reliability 98.3%). Data collectors arrived at the preschool at times arranged with the teacher or administrator, and conducted assessments of the participating child's self-regulation via the *HTKS*. Data collectors removed the child from his or her classroom and followed a script (Appendix F) inviting him or her to play a game and obtaining verbal assent. If the child did not have a completed packet or parental consent, refused to participate, or elected to withdraw, the next child in the randomly generated order was offered to participate following the same procedures.

Completion of the *HTKS* task occurred in a separate room or quiet space free from distractions. Some sites required school staff to be present for the assessment; in these situations the staff sat quietly behind the student out of their sightline to minimize distraction during the task. Following completion of the self-regulation assessment, children selected their choice of a small prize for their time (valued less than or equal to $1) and were allowed to return to their classroom. Data collectors provided packets of questionnaires to the participating child's teacher and asked her to complete and return the packets to the researcher in self-addressed and stamped envelopes. Data collectors also obtained teachers' written consent at this time if it had not already been received. Teachers received a donation of age-appropriate books for

their classrooms as compensation for their time (valued at approximately $10-12). Teachers were contacted after two weeks to remind them to complete and return packets. Teachers who had not returned the packets after four weeks were contacted again, and offered a scheduled time at which the researcher could pick up the completed packets from the teacher's school. Three teachers elected to withdraw from the study and did not complete rating forms.

Data Analyses

All data analyses were completed using SAS Version 9.4, University Edition (SAS Institute, 2018). Data analyses were conducted in three waves. In the first wave, *Preliminary Analyses,* descriptive statistics for the study's variables were identified to ensure data met necessary assumptions for analyses. Group mean differences using Analysis of Variance (ANOVA) procedures and Pearson's correlation coefficients were obtained for proposed covariates and the models' criterion variables to determine whether to include covariates in the final models.

The second wave of analyses comprised *Regression Analyses.* First, multiple linear regression was examined using SAS PROC REG procedures to explore the predictive relationship between the *HTKS* task and the three behavioral measures (*C-TRF, EC BEH-T,* or *SSIS* scales) as continuous variables. Next, the predictive validity of the *HTKS* for clinically elevated externalizing problem behaviors (i.e., *T*-scores $\geq$ 60) was examined for the included scales of the *C-TRF* and *EC BEH-T* via logistic regression analyses to explore the classification probabilities of the *HTKS* task. The *SSIS* did not produce any clinically elevated *T*-scores and was excluded from second wave analyses. A binomial logistic regression model was implemented using SAS

PROC LOGISTIC. Total score on the *HTKS* served as the predictor variable, and separate analyses were run for each measure of externalizing problem behavior (i.e., two models were run using the externalizing problem behavior scales from the *C-TRF* and *EC BEH-T* as the respective criterion variables). The model for the *C-TRF* included previous preschool experience as a covariate.

Descriptions of the statistical models follow. To promote clarity, elements of the models have been simplified. That is, covariates are listed as general child covariates, but specific covariates were added to the model at the time of analyses. As well, independent models were run for each measure of externalizing problem behaviors (i.e., problem behavior scales from the *C-TRF, EC BEH-T,* and *SSIS*), but only a single model is presented using a general "behaviors" variable. This general variable was replaced with each respective model's rating form measure of behavior problems.

The multiple linear regression equation was:

$$\widehat{\text{Disruptive Behavior Problems}}_i = b_0 + b_1(\text{HTKS}_i) + b_2(COVS_i) + e_i$$

In this model, $\widehat{\text{Disruptive Behavior Problems}}_i$ is the predicted *T*-score of that model's behavior scale. For the model's predictors, b_0 represents the model's intercept and $b_1(HTKS_i)$ represents the main effect of children's self-regulation (measured by total score on the *HTKS* task). The next regression coefficient represents the fixed covariate effects, where $b_2(COVS_i)$ indicates included covariates (i.e., child age, preschool experience, ethnicity/race minority status, and maternal education level). Finally, e_i represents the net residual term.

The binomial logistic regression equation was:

$$\widehat{\text{Logit Behavior}}_i = b_0 + b_1(HTKS_i) + b_2(COVS_i) + e_i$$

In this model, $\widehat{\text{Logit Behavior}}_i$ is the predicted logit (i.e., log of the odds = $\ln\left[\frac{\hat{\Pi}_i}{1-\hat{\Pi}_i}\right]$, where $\hat{\Pi}_i$ = the probability of placement in group) of a child being rated as demonstrating clinically elevated externalizing behavior problems (*T*-score ≥ 60) for child *i.* It should be noted that the use of a logit in the equation provides for a traditionally formatted general linear model; however, the interpretation of the final results report the model's odds (odds = e^{logit}) and odds ratio (comparing change in the odds for a one-unit increase in the predictor, such that odds ratio from values 2 to 3 $= \frac{\text{odds}(3)}{\text{odds}(2)} = \frac{e^{b_0}e^{3b_1}}{e^{b_0}e^{2b_1}} = e^{b_1}$) which is more conceptually aligned with the study's research question. That is, this conversion allowed for results to include an interpretation of the change in the odds of a child being in the clinically elevated behavior group for each unit increase in *HTKS* score. The odds ratio has a range between 0 and infinity, with an odds ratio of 1 indicating no association between the criterion and predictor. Therefore, as the value decreases to zero or increases to infinity away from 1, the association is said to be more powerful.

In the remainder of the model, b_0 represents the model's intercept and $b_1(HTKS_i)$ represents the main effect of children's self-regulation (measured by the *HTKS* task). The next regression coefficient represents the fixed covariate effects, where $b_2(COVS_i)$ indicates covariates (i.e., child age, preschool experience, ethnicity/race minority status, and maternal education level). Finally, e_i represents the net residual term.

The resulting model from the logistic regression was used to develop a classification table, which helps visualize the percentage of true and false predictions. This allows a comparison of the model's specificity (i.e., true negative rate) against its sensitivity (i.e., true positive rate) to produce a total accuracy score, using the equation $Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives}$. Table 2 demonstrates an example classification table.

Table 2

Example Classification table.

		Predicted	
		Clinically Elevated	Normative
Actual	Clinically Elevated	True Positive	False Negative
	Normative	False Positive	True Negative

Further analysis of the sensitivity and specificity of the *HTKS* task utilized ROC (Receiver Operating Characteristics) curves (Swets, Dwawes, & Monahan, 2000) and comparison of the area-under-the-ROC-curve (AUC) values for *C-TRF* and *EC BEH-T* scales. ROC curve analyses operate by plotting the sensitivity (i.e., likelihood of a type I error) against the specificity (i.e., likelihood of a type II error) for given values (cutpoints) of the predictive measure (i.e., obtained scores on the *HTKS* task). The resulting graph presents the rate of true positives on the *y*-axis, and the rate of false positives (1 – specificity) on the *x*-axis (Figure 1).

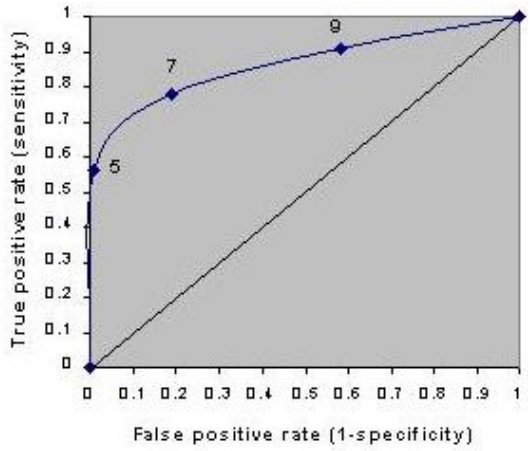

Figure 1. Sample ROC Curve Graph. By T. G. Tape, 2017. *Plotting and interpreting an ROC curve.* Retrieved from http://gim.unmc.edu/dxtests/ROC2.htm

The straight line which bisects the plot in Figure 1 is used for comparison during AUC analyses. If the sensitivity and specificity of the model were identical (i.e., no better than chance), the ROC curve would lie directly on the straight line in the plot. The AUC value is obtained by calculating the area of the plot below the ROC curve. As a test improves in its ability to discriminate outcomes, the AUC values increase. The AUC values range from 0.5 (no better than chance) to 1.0 (100% sensitive and 100% specific). General "rules-of-thumb" for AUC values may then be used to assess the quality of discrimination provided by the predictor variable (Hosmer, Lemeshous, & Sturdivant, 2013). Those values may be seen in Figure 2.

	AUC = 0.5	No discrimination
If…	$0.5 < AUC < 0.7$	Poor discrimination
	$0.7 \leq AUC < 0.8$	Acceptable discrimination
	$0.8 \leq AUC < 0.9$	Excellent discrimination
	$AUC \geq 0.9$	Outstanding discrimination

Figure 2. Guidelines for interpreting AUC values. Adapted from *Applied Logistic Regression (3rd Ed.)* by D. W. Hosmer, S. Lemeshow, & R. X. Sturdivant, 2013. Copyright 2013 by John Wiley and Sons.

CHAPTER 4: RESULTS

Preliminary Analyses

Descriptive statistics (means, standard deviations, skewness, and kurtosis) for study variables were obtained and are reported in Table 3. *HTKS* Total refers to the total obtained score for each participant on the *HTKS* task. *HTKS* Time is the time (in seconds) participants took to complete the task. *T*-scores for the Externalizing Problems Scale of the CBCL *C-TRF,* and Defiant/Aggressive Behaviors Scale of the Conner's *EC-BEH-T* were obtained using the respective measures' scoring software. The final measure of disruptive behaviors was the Problem Behaviors Scale of the *SSIS*. The *SSIS* produces standard scores (M = 100, SD = 15), which were converted into T-Scores using the formula $T = \frac{2}{3}SS - 16$ to facilitate comparisons between measures[1]. Descriptive statistics for the Problem Behaviors Scale *T*-Score of the *SSIS* are also reported in Table 3.

Table 3

Descriptive Statistics

Variable	*N*	*M*	*SD*	Range	Skewness	Kurtosis
HTKS Total Score	24	22.08	9.16	0–37	-0.47	-0.58
HTKS Time (in seconds)	24	306.12	41.25	235–401	0.45	-0.43
C-TRF Externalizing Problems	21	54.00	10.45	36–69	-0.39	-1.31
EC BEH-T Defiant/Aggressive Behaviors	21	57.48	14.09	41–87	0.44	-1.23
SSIS Problem Behaviors	21	45.75	5.16	40–54	0.23	-1.71

Note. HTKS possible range = 0–40. Scores reported for *C-TRF*, *EC BEH-T*, and *SSIS* are *T*-scores (i.e., $\bar{x}$ = 50, *SD*=10)

In this sample, all of the measures' data had a skewness greater than -0.5 and less than 0.5, indicating the distribution is approximately symmetric (Bulmer, 1979). Further, George and Mallery (2010) argue that normality can be assumed when kurtosis and skewness fall between -2 and 2, indicating this sample meets this assumption and thus no transformations were necessary.

Controlling for group differences and covariates. Group means were compared via Analysis of Variance (ANOVA) procedures in PROC GLM for the proposed categorical covariates (Gender, Ethnic Minority Status, Parents' Combined Income Range, Maternal Education, and Previous Preschool Experience) for each of the three criterion variables and *HTKS*. Significant *F* scores suggested the presence of group differences in mean disruptive behavior scores or *HTKS* performance, respectively. Appendix G compiles the tables from these analyses and provides the group means and *F* statistic for each comparison. Only previous preschool experience possessed significant group differences for *C-TRF*, $F(4,16) = 3.12$, $p = 0.04$ (Table 4). None of the measured covariates presented with significant group mean differences for either *EC BEH-T* or *SSIS* scales or the *HTKS* (Appendix G). As a result, preschool experience was included as a covariate in the *C-TRF* models but the remaining models did not include covariates.

Table 4

ANOVA for Previous Preschool Experience on C-TRF Externalizing Problems

	Sum of Squares	df	Mean Square	F
Model	957.57	4	239.39	3.12*
Error	1228.43	16	76.78	
Total	2186.00	20		

*$p < 0.05$

Initial bivariate analyses explored the relationship between the study's variables. Pearson's correlation coefficients were obtained for the continuous variables included in this study (Table 5).

Table 5

Correlation Matrix for Continuous Variables

Variables	1.	2.	3.	4.	5.
1. Head Toes Knees Shoulders Total Score (*N* = 24)	-				
2. SSIS Problem Behaviors T-Score (*N* = 21)	0.464*	-			
3. C-TRF Externalizing Problems T-Score (*N* = 21)	0.128	0.813**	-		
4. EC BEH-T Defiant/Aggressive Behaviors T-Score (*N* = 21)	0.417	0.965**	0.835**	-	
5. Child age (months; (*N* = 23)	0.238	-0.002	-0.205	0.104	-

*p<.05; **p<.001

Correlation coefficients revealed a significant moderately-sized positive relationship between *HTKS* performance and the *SSIS* problem behaviors *T*-score (r = 0.464, p = 0.03), such that children with higher performance on the *HTKS* task tended to have higher (i.e., more severe) problem behaviors. A similar pattern emerged for the *EC BEH-T* task and *HTKS* (r = 0.417, p = 0.06); however, this correlation was non-significant. The *C-TRF's* correlation with *HTKS* was non-significant and relatively small (r = 0.128, p = 0.58). As a result, none of the three measures produced a significant linear relationship with the *HTKS* task in the hypothesized direction. Inspection of scatter plots did not identify any outliers interfering with observed relationships. However, analyses were continued to further explore whether this was

attributable to covariate effects or if the measures operated better as predictors when dichotomized (i.e., as in clinical practice when these forms are used to inform diagnostic decisions).

Regression Analyses

The second wave of data analyses employed multiple linear regression and logistic regression methodology. These analyses sought to answer the study's research questions:

1. *"Does temperamental self-regulation predict externalizing behaviors?"*
2. *"How well does a performance measure of temperamental self-regulation differentially predict preschoolers who demonstrate clinically elevated externalizing behaviors from those preschoolers who do not?"*

Multiple linear regression analyses explored the relationship between criterion (behavior rating scales) and predictor (self-regulation and demographic characteristics) variables. Obtained T-scores for the *SSIS*, *C-TRF*, and *EC BEH-T* measures were recoded into dichotomous variables based on the cut point (0 = T-score < 60; 1 = T-score $\geq$ 60). No scores on the *SSIS* fell in the clinically elevated range (i.e., T-score $\geq$ 60); therefore the *SSIS* could not be dichotomized and analyses were discontinued on the *SSIS*. Both the *C-TRF* and *EC BEH-T* identified eight children who met criterion. Five children's scores were clinically elevated on both measures, while six children had clinically elevated scores in either the *C-TRF* or *EC BEH-T*. Five logistic regression analyses then explored the *HTKS'* performance predicting children with clinically elevated scores on the *C-TRF* and *EC BEH-T* from those without.

Predicting Problem Behaviors Scale of the SSIS

A multiple linear regression model was fit to examine the relationship between *HTKS* and the Problem Behaviors scale of the *SSIS* as well as covariates. Table 6 summarizes the results. The multiple regression model with all of the predictors produced $R^2 = 0.388$, $F(5,15) = 1.90$, $p > .15$. As can be seen in Table 6, among the predictors only the *HTKS* had a significant regression weight. After controlling for the other variables in the model, the *HTKS* had a significant, positive weight, such that for each one unit increase in *HTKS* the Problem Behavior scale *T*-score of the *SSIS* increases 0.26. This is consistent with the relationship identified in correlation analyses (Table 5), indicating no suppressor effect is present. These findings are in the opposite direction of the hypothesized relationship between *HTKS* and *SSIS*.

Table 6

Results from regression analysis for SSIS on HTKS and covariates (N = 21)

Predictor	β	SE	*p*
Intercept	41.37*	10.63	0.001
HTKS	0.26*	0.12	0.038
Age	-0.07	0.18	0.696
Gender	-1.56	2.32	0.512
Ethnic Minority	2.08	0.85	0.410
Previous Preschool	1.22	0.734	0.116

**p<.05*

Predicting Externalizing Behaviors Scale of the CBCL (C-TRF)

Multiple regression. A multiple linear regression model was fit to examine the relationship between *HTKS* and the Externalizing Behaviors scale of the *C-TRF* as well as covariates. Table 7 summarizes the results. The multiple regression model with all of the predictors was nonsignificant; $R^2 = 0.279$, $F(5,15) = 1.16$, $p = 0.374$. As can be seen in Table 7, none of the predictors had significant regression weights.

Table 7

Results from regression analysis for C-TRF on HTKS and covariates (N = 21)

Predictor	β	SE	*p*
Intercept	59.93*	23.38	0.022
HTKS	0.14	0.26	0.594
Age	-0.298	0.401	0.469
Gender	1.249	5.11	0.810
Ethnic Minority	4.07	5.41	0.464
Previous Preschool	3.26	1.61	0.061

**p<.05*

Logistic regression. A two-predictor logistic model was fit to the data to test the hypothesis regarding the relationship between a child's clinically elevated scores on the Externalizing Problems subscale of the *C-TRF* and his or her performance on the HTKS task and previous preschool experience. The results showed the log odds of a child being rated with clinically elevated behavior problems was not related to performance on the *HTKS* ($p > .05$) nor previous preschool experience ($p > .05$, Table 8). Using Chen, Cohen, and Chen's (2009) standards for interpreting the effect size odds ratios (i.e., Cohen's $d < 0.2$ when OR <1.5, and Cohen's $d > 0.8$ when OR > 5; Chen et al., 2009), this model produced odds ratios with small (*HTKS, OR = 0.97)* to medium (Previous experience, *OR = 2.06)* effect sizes[2]. Interpretation of these odds

ratios suggests that, holding the other predictor equal, the odds of a student being identified with clinically elevated behavioral concerns would decrease by approximately 3% for each one-unit increase on the *HTKS.* The odds ratio of 2.06 for preschool experience suggests that in this sample, the odds of students being identified with clinically elevated behavioral concerns on the *C-TRF* approximately doubles for each additional year of center-based preschool experience. None of the model's χ^2 tests of fit (Likelihood ratio $\chi^2 = 4.4$, score test $\chi^2 = 4.05$, Wald test $\chi^2 = 3.49$; all $p > .05$) were significant, suggesting the model is no different from the null model (i.e., without predictors) in predicting elevated behavioral symptoms status on the *C-TRF*. The inferential goodness-of-fit test used is the Hosmer–Lemeshow (H–L) test, which yielded a $\chi^2(7)$ of 7.26 and was insignificant ($p > .05$), suggesting that the model was fit to the data well.

Table 8

Logistic regression analyses of elevated C-TRF scores (N = 21)

Predictor	β	*SE* β	Wald's χ^2	*df*	*p*	e^{β}
Intercept	-1.42	1.53	0.86	1	0.35	0.24
Head Toes Knees Shoulders Total Score	-0.03	0.06	0.34	1	0.56	0.97
Previous Center-based Preschool Experience	0.72	0.39	3.41	1	0.07	2.06
Test			χ^2	*df*	*p*	
Overall model evaluation						
Likelihood ratio test			4.40	2	0.111	
Score test			4.05	2	0.132	
Wald test			3.49	2	0.175	
Goodness-of-fit test						
Hosmer & Lemeshow			7.26	7	0.402	

A classification table of *HTKS* predicting *C-TRF* (Table 9) revealed probability level for the classification table was set to 0.5; that is, the model's performance in identifying clinically elevated behavior concerns was compared against chance (50%).

Table 9

Classification table of HTKS predicting C-TRF Externalizing Problems scale

		Predicted	
		Clinically Elevated	Normative
Actual	Clinically Elevated	2	6
	Normative	3	10

The model's accuracy can be computed using the following formula:

$$\frac{True\ Positives+True\ Negatives}{True\ Positives+True\ Negatives+False\ Positives+False\ Negatives} = \frac{2+10}{2+10+6+3} = \frac{12}{21} = 57.1\%$$

accuracy

Thus, the logistic regression model does not appear to provide sufficient accuracy above and beyond chance (expected 50%). Further, this model only produced two true positives (25% sensitivity), suggesting a poor fit for its proposed use as screening tool.

Predicting Defiant/Aggressive Behavior Scale of the Conners (EC BEH-T)

Regression analyses. A multiple linear regression model was fit to examine the relationship between *HTKS* and the Defiant/Aggressive Behavior scale of the *EC BEH-T* as well as covariates. Table 10 summarizes the results. The multiple regression model with all of the predictors produced $R^2 = 0.328$, $F(5,15) = 1.47$, $p = 0.258$. As can be seen in Table 10, none of the predictors had significant regression weights.

Table 10

Results from regression analysis for EC BEH-T on HTKS and covariates (N = 21)

Predictor	β	SE	*P*
Intercept	29.42	30.42	0.349
HTKS	0.59	0.33	0.10
Age	0.12	0.52	0.815
Gender	-3.30	6.65	0.627
Ethnic Minority	3.90	7.04	0.588
Previous Preschool	3.58	2.10	0.109

**p<.05*

Logistic regression. A one-predictor logistic model was fit to the data to test the hypothesis regarding the relationship between a child's clinically elevated scores on the Defiant/Aggressive Behaviors subscale of the *EC BEH-T* and his or her performance on the *HTKS* task (Table 11). The results indicated that the log odds of a child being rated with clinically elevated behavior problems was not related to performance on the *HTKS* ($p > .05$). This model produced an odds ratio for the *HTKS* ($OR = 1.12$) a small effect size (Chen et al., 2009). Interpretation of this odds ratio suggests that for each additional point earned on the *HTKS* task, the odds of a student being identified with clinically elevated behavioral concerns would increase by approximately 12%. One of the model's χ^2 tests of fit (Likelihood ratio $\chi^2 = 4.03$; $p = .04$) was significant. However, remaining tests of fit (score test $\chi^2 = 3.50$, Wald test $\chi^2 = 2.94$; both $p > .05$) were not significant. When the tests of fit do not yield similar conclusions, Menard (1995) recommended to rely upon the likelihood ratio and score tests only. This standard still produces an ambiguous interpretation of the model's fit, so a definitive statement about whether the model worked better than chance in predicting elevated behavioral symptoms status on the *EC-BEH-T* is not possible.

Because the predictor in the model was not significant and thus no meaningful change was predicted via the model, the results of the fit tests indicating no difference from the null appear to be more representative of the current model. The (H–L) test yielded a $\chi^2(5)$ of 5.62 and was insignificant ($p > .05$), suggesting that the model was fit to the data well.

Table 11

Logistic Regression Analyses of elevated EC BEH-T scores

Predictor	β	*SE* β	Wald's χ^2	*df*	*p*	e^{β}
Intercept	-3.33	1.82	3.36	1	0.07	0.036
Head Toes Knees Shoulders Total Score	0.12	0.07	2.94	1	0.09	1.13
Test			χ^2	*df*	*p*	
Overall model evaluation						
Likelihood ratio test			4.03	1	0.045	
Score test			3.50	1	0.061	
Wald test			2.94	1	0.087	
Goodness-of-fit test						
Hosmer & Lemeshow			5.62	5	0.345	

A classification table was also built for this model (Table 12). The model only produced three true positives (37.5% sensitivity). The overall accuracy of 57.1% provided evidence this model did not discriminate children with clinically elevated behavior symptoms much better than chance.

Table 12

Classification table of HTKS predicting EC BEH-T Defiant/Aggressive scale

		Predicted	
		Clinically Elevated	Normative
Actual	Clinically Elevated	3	5
	Normative	4	9

Follow-up/Exploratory Analyses

Both *C-TRF* and *EC BEH-T* measures produced eight clinically elevated scores and 13 normative scores. The measures did not identify the same eight children; thus, the selected measures also did not provide identical sensitivity and specificity regarding the overarching construct of externalizing behavior problems. This is further supported by the absence of any clinically elevated scores among children rated using the SSIS. *Post hoc* exploratory analyses to further elucidate the second research question were conducted. These included exploring the resulting ROC (Receiver Operating Characteristics) curves (Swets, Dawes, & Monahan, 2000) and comparing the resulting area-under-the-ROC-curve (AUC) values. Although the logistic regressions did not produce significant models, these additional analyses were conducted for the purpose of further exploring the potential of performance on the *HTKS* as predictive of behavioral problems. Additionally, initial cost-benefit analyses were explored for future studies implementing *HTKS* in school settings.

ROC curve analyses. ROC curve analyses operate by plotting the sensitivity (i.e., likelihood of a type I error) against the specificity (i.e., likelihood of a type II error) for given values (cutpoints) of the predictive measure (i.e., obtained scores on the *HTKS* task). ROC curve graphs were generated for the *C-TRF* (Figure 3) and *EC BEH-T* (Figure 4). The AUC for the *C-TRF* model was 0.52, suggesting poor

discrimination (i.e., comparable to chance). The *EC BEH-T*, however, produced an AUC of 0.73, suggesting acceptable discrimination. Visual inspection revealed a cut score of 14 on the *HTKS* appeared to demonstrate a sufficient balance of sensitivity (.75) against specificity (.46). However, it should be repeated that these findings were in the opposite direction as hypothesized, such that students with higher scores on the *HTKS* are more likely to present with clinically elevated behaviors. Thus, these findings warrant further investigation in future research, and using the *HTKS* with a cut score for identification is not recommended at this time.

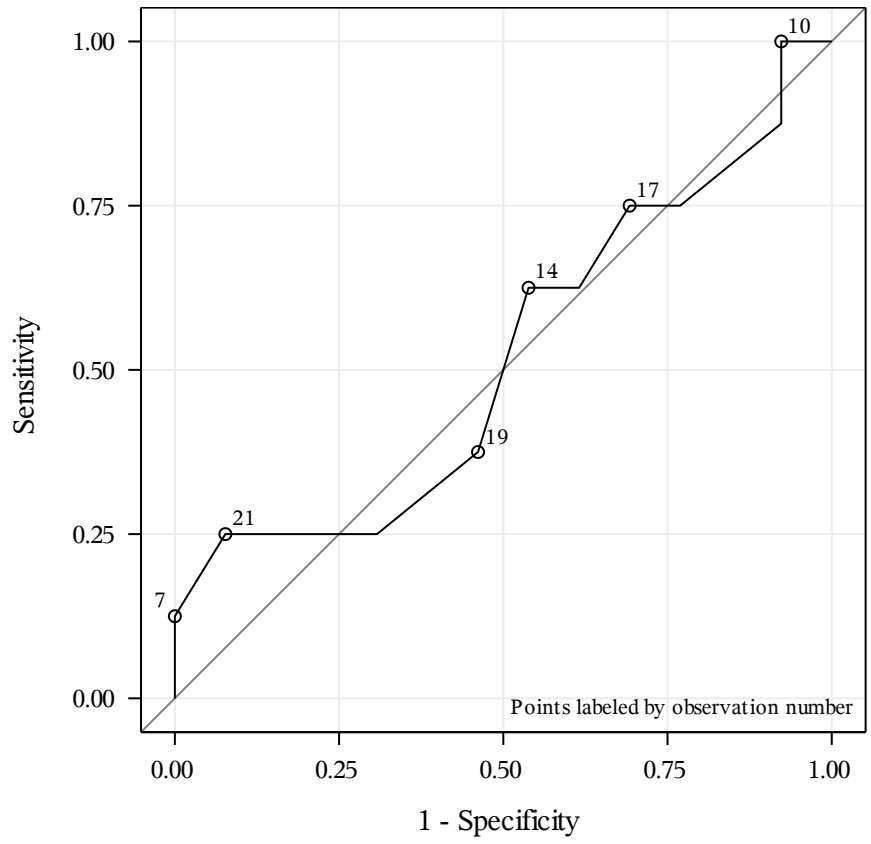

Figure 3. ROC Curve for *HTKS* predicting *C-TRF*. AUC = 0.5192.

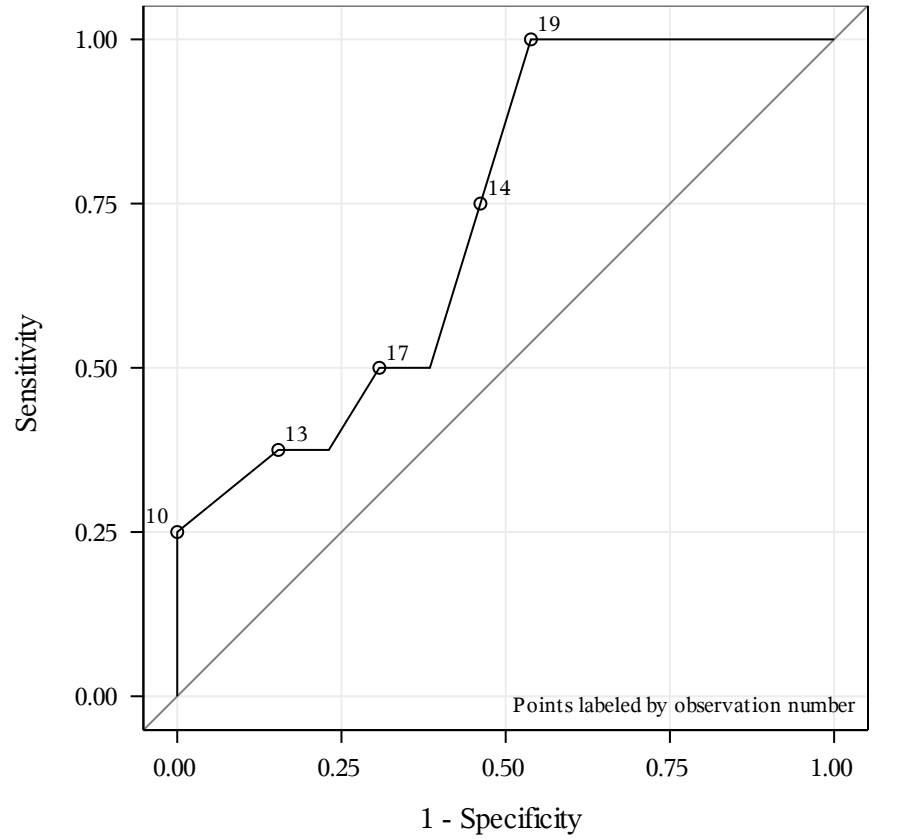

Figure 4. ROC Curve for *HTKS* predicting *EC BEH-T*. AUC = 0.7360.

Cost-benefit analyses. This study aimed in part to explore the utility of the *HTKS* in a universal screening context for preschoolers. Administration of the task took an average of 306 seconds (range = 235–401 seconds). This replicated previous administration times in previous research (e.g., McClelland et al., 2014) which suggested the *HTKS* requires approximately five minutes to administer, on average. It should be noted, however, that the third quartile of the distribution occurred at 326.8 seconds, suggesting a quarter of participants required approximately five and one-half minutes or longer to complete the task. Thus, full screening of a classroom of 20 preschoolers' self-regulation should take approximately 100 minutes of assessment

time to complete. The design of the present study (i.e., cross-sectional sampling) did not allow comparison of the *HTKS* task's utility as a screener above and beyond existing screening measures (e.g., rating forms; "wait and see" approach). Rather, the study only assessed whether the *HTKS* could identify children currently displaying externalizing behavior problems. However, the current study's inability to reject the null hypotheses provides preliminary evidence (with the absence of longitudinal data or a confirmatory diagnosis) that empirically validated behavioral screeners could outperform the *HTKS* in their ability to identify children most at risk for behavioral concerns.

CHAPTER 5: DISCUSSION

Main Findings

This study had two primary research questions. The first question asked whether temperamental self-regulation (measured by the *HTKS* task) predicted externalizing behaviors among preschoolers (i.e., *T*-scores on select externalizing behavior scales of the *C-TRF, EC BEH-T,* and *SSIS*). Pearson correlation coefficients for *SSIS* and *EC BEH-T* scores identified a moderate, positive relationship between their respective scores and the *HTKS*. This relationship was in the opposite direction than hypothesized, such that children with higher self-regulation as measured by the *HTKS* appeared to tend to have higher (i.e., more problematic) behavior concerns on the *SSIS* and *EC BEH-T* measures. Multiple linear regression analyses sought to further elucidate the predictive relationship between externalizing behaviors and self-regulation by fitting the variables in a predictive linear model. Regression analyses across the three behavioral measures did not produce any significant models for the prediction of behavior concerns. Among the three models, only one (*SSIS*) produced a significant predictor. In that model the *SSIS* Problem Behaviors scale was significantly predicted by the *HTKS* task; however, it was in the opposite direction as hypothesized such that higher scores on the *HTKS* predicted increases in problem behaviors.

The second question sought to identify how well a performance measure of temperamental self-regulation differentially predicts clinically elevated behaviors among preschoolers. Although the initial analyses did not support the hypothesized predictive relationship between self-regulation and behavior concerns, binomial logistic regression analyses explored whether the relationship between the measures

used would be more useful if the criterion variable was dichotomized (i.e., clinically elevated vs. normative behavior problems). The *SSIS* did not produce any significantly elevated *T*-scores and was excluded from analyses. Logistic regression and follow-up analyses explored the performance of the *HTKS* task in predicting clinically elevated behavior problems for the *C-TRF* and *EC BEH-T*. Based upon extant literature, it was hypothesized that children with higher scores on the *HTKS* (i.e., more regulated) would be less likely to be rated with clinically elevated behavior problems by their teachers. However, neither of the models identified significant changes in the odds of a student demonstrating clinically elevated behavior concerns based on changes to *HTKS.* Follow-up visual inspection of classification tables and ROC curve plots indicated that children's self-regulatory abilities (as indicated by total scores on the *HTKS*) did not appear to accurately identify children at-risk for developing an externalizing behavior disorder (i.e., children with clinically elevated scores on standardized rating measures of behavior disorders), contrary to the study's hypotheses. Although the ROC curve and AUC value for the *EC BEH-T* did demonstrate qualitatively "acceptable" discrimination, the direction of the relationship was opposite this study's hypotheses (i.e., children with higher self-regulation were more accurately classified with clinically elevated behavior concerns). The following section further explores these unexpected results and seeks to provide context for their interpretation.

Unexpected Results and Interpretation

This study sought to provide preliminary evidence toward a robust trajectory of future research using self-regulation to screen for behavioral risk. However, findings

did not allow for the rejection of the null hypotheses. This study's findings suggest that a single behavioral task of self-regulation such as the *HTKS* may not be directly predictive of children's externalizing behavior disorders as rated at the time of assessment. It is important to note, however, that this study's failure to reject its null hypotheses does not disprove a negative predictive relationship between self-regulation and externalizing behavior disorders. Rather, this study was unable to find sufficient statistical evidence to support such a relationship.

The temperamental construct of self-regulation is expected to remain stable; however, any measure of temperament should be considered along with the child's environmental context (Berdan, Keane, & Calkins, 2008). This study's findings, although surprising, do provide evidence that the relationship between self-regulation and disruptive behaviors may be more complicated than initially conceptualized. For this study, a single child was pulled from the classroom and began assessment after brief rapport-building. The administration of the *HTKS* features a highly engaging activity delivered in a reinforcing one-on-one setting. It is possible this mode of administration was differentially reinforcing for students whose disruptive behaviors are functionally maintained by attention in classroom settings. Such conditions could confound the *HTKS'* ability to predict children's disruptive behaviors. That is, students for whom individual attention is most rewarding may have been especially motivated to perform well on the *HTKS*. Further, the inclusion of a small prize for completing the task could have modulated the students' motivation for participation. In fact, utilizing a measure such as the *HTKS* which purports to assess temperamental self-regulation (i.e., effortful control or "hot" self-regulation; Willoughby et al., 2011) left

participants susceptible to effects on their motivation in a way tasks assessing "cool" self-regulation may not (see Denham et al., 2015). The hypothesized relationship between effortful control and externalized behavior problems was supported in the literature (e.g., Espy et al., 2011); however, this study was unable to capture that relationship.

This is especially meaningful in the context of an MTSS/PBIS framework. The findings of this study suggest that simply assessing self-regulation using the *HTKS* is not likely to produce sufficient information to determine which children are most likely to present with elevated behavioral symptoms. These null findings provide an interesting addition to the literature; that is, temperamental self-regulation may not be the construct driving children's behavioral concerns. Indeed, the suggestion that self-regulation's predictive ability of behavioral concerns may vary by mode of assessment (and its capacity for delivering reinforcing attention) provides ample opportunity for additional research questions and designs.

Future Directions

The results of this study do not support the hypotheses that the *HTKS* would (a) negatively predict externalizing behavior problems; and (b) differentially predict children with clinically elevated behavior problems from those without, such that children with lower scores had increased risk of clinically elevated symptoms. The findings raised several questions which merit further exploration.

First, despite previous research establishing correlations between the *HTKS* and behavior problems (e.g., Montroy et al., 2014), findings from this study did not provide evidence for the task alone as a unique predictor of clinically elevated

externalizing behavior. Future studies may wish to explore task paradigms designed to assess specific domains of self-regulation (e.g., inhibition paradigms; Carlson & Moses, 2001) to screen for behavioral concerns, in contrast to the current study's attempt to screen via the overarching construct of self-regulation. Future studies may also benefit by exploring different modes of assessment (e.g., group administrations vs. individual; computer-based vs. in-person). It is likely school-based universal screening would not include prizes for the children as part of a school-based assessment, so another empirical question might examine the extent to which completing tasks for a prize activates the "hot" versus "cool" systems of self-regulation in screening contexts.

Next, the correlational analyses in this study suggested a positive linear relationship between the *HTKS* and measures of externalizing problem behaviors. This was unexpected based on extant literature including these constructs. Future research may benefit from further exploring the relationship between *HTKS* and externalizing behaviors, including whether the demographic characteristics of this study's homogeneous sample comprise conditions in which the relationship is moderated. For example, it is certainly possible the relationship between self-regulation and externalizing behaviors could be fundamentally different between children from high- and low-socioeconomic status households (e.g., extant evidence has shown children from high-income households utilize neural systems while engaging in self-regulatory performance tasks differently from children from lower-income households; Hackman & Farah, 2009). Future research may wish to explore longitudinally whether the relationship is causal or perhaps bidirectional (e.g., perhaps children with more

externalizing behaviors receive more practice regulating those impulses and thus improve their regulatory capacity).

Future studies may also benefit from exploring the utility of the *HTKS* (or a similar measure of self-regulation) within a single system (e.g., all students in a single preschool location). This would allow comparison of students' scores relative to their peers in the same context. The current study sought to control for teacher effects by randomly selecting one child per teacher to participate; however, in the context of a MTSS service delivery framework, students ideally receive increasingly intensive intervention services based on their screening data relative to base rates in their own school's system (see Kilgus & Eklund, 2016).

The results of the current study did not support the use of self-regulation for screening children's risk for behavior problems. However, it should be noted that the importance of self-regulation as a construct in school readiness is not called into question by these findings. A robust literature supports self-regulation's role in students' academic (e.g., McClelland et al., 2014) and socio-emotional (e.g., Lonigan et al., 2017) readiness. Even if the *HTKS* task is unable to provide sufficient sensitivity and specificity to effectively screen for preschoolers' clinically elevated behavioral symptoms, the task has been validated for use as a measure of a construct which permeates academic functioning. Therefore, future research may benefit instead from exploring self-regulation as a target behavior for assessment and intervention, rather than as a substrate for other constructs (i.e., elevated behavioral problems, as in the current study).

Study Limitations

Results of the current study should be interpreted with caution due to several limitations that may affect validity. First, this study presented with threats to statistical conclusion validity. The study was cross-sectional in design and lacked experimental control. As a result, causal inference cannot be made between the criterion and predictor variables. Further, the sample size in this study (*N = 24*) may have been insufficiently large to detect significant relationships between variables. The current study achieved, but did not exceed, the necessary sample size suggested by *a priori* power analyses. As a result, deviations from the conditions of the data used for power analyses increase the risk of the current study failing to detect true effect sizes. For example, it is possible the *HTKS* may have a different relationship with the criterion variables than the Fish/Shark task used for power analyses. The Fish/Shark task used in power analyses was administered in a lab-based setting on a monitor without child incentives (i.e., conditions exemplifying "cool" self-regulation), whereas the current study administered the *HTKS* under increased "hot" conditions (i.e., face-to-face in the child's school and delivering a prize for the child's completion). Finally, the use of a specified "go/no-go" task paradigm in the power analyses may have confounded the current study's results as the *HTKS* and Fish/Shark measures appear to be assessing slightly different constructs.

Second, characteristics of the sample in the current study may have threatened the internal validity of the study. The sample was homogeneous, presenting as disproportionately affluent (87.5% of combined family incomes greater than $100,000), white (87.5%), and male (69.6%). Participants were also geographically

restricted to the Lincoln and Omaha metropolitan areas attending private or for-profit preschools. No children attending public preschools participated in this study. Participating children were randomly selected in an attempt to minimize selection bias in the study. However, children available for selection were drawn from families who signed and completed packets to classrooms which also agreed to participate. The number of packets returned varied widely by classroom, and there is a risk that parents who elected to participate presented with differences from those electing not to participate (e.g., demographic characteristics, child's behavioral history). Further, it is possible the small sample size and homogeneous sample failed to obtain sufficient diversity among behavior problems identified by the measures.

A third limitation involves the interpretation of the measures used in the context of this study. It is possible the measures selected to assess the construct of "externalizing behavior problems" may have been subject to nonrandom error. The measures were provided to teachers to rate randomly assigned individual students. However, teachers were not randomly selected to participate. As a result, there may have been teacher-level factors which impacted their ratings of a given student. Further, this represents a cross-sectional sampling of student behaviors (i.e., the teachers' perspective of student functioning at the time of the assessment). Longitudinal assessment of student behaviors (and tracking actual referrals for behavioral concerns) could have provided more comprehensive information of underlying behavioral concerns. That is, this study merely was unable to reject the null hypothesis to determine whether *HTKS* is an effective predictor of teacher reported

behavioral concerns at the time of assessment, rather than later development of behavior disorders.

Conclusion

The purpose of this study was to explore whether the *HTKS* task can accurately predict preschoolers who demonstrate clinically elevated behavior problems from those who do not. This study sought to establish the theoretical underpinnings for a program of research establishing the validity of universally screening self-regulation within a multi-tiered systems of support framework. Descriptive analyses revealed no group differences among proposed covariates, with the exception of previous preschool experience and performance on the *C-TRF* Externalizing Problems scale. The sample obtained in this study was overwhelmingly affluent, white, and male. Results of regression analyses did not support the study hypotheses, and none of the regression models were significant. However, some evidence was found for small-to-moderate correlations between *HTKS* scores and behavioral concerns. These relationships were in the opposite direction as hypothesized.

Several limitations impede the interpretation of these findings; however, and future research must address these limitations before a definitive conclusion may be drawn regarding the utility of implementing universal screening for self-regulation. Future research is necessary to clarify correlational and longitudinal relationships between behavior problems and self-regulation. Aside from methodological limitations in the present study as described above, a future study should determine whether the positive correlation of self-regulation with behavior problems can be found with other measures of self-regulation, or whether this is an anomalous event under the

conditions of this study. For example, it is possible that self-regulation differentially predicts self-regulation across development and is unable to serve as a predictor at the concurrent time frame of this study. That is, perhaps behavior problems for students with low self-regulation appear later in their schooling. Further, the *HTKS* task itself may not capture the construct of self-regulation in a manner that is useful for predicting behavior concerns. Indeed, it is possible that either the study's measure of self-regulation (i.e., *HTKS*) or the construct of self-regulation itself may not be related to behavior as predicted, despite previous suggestions in extant literature. Further exploration of these findings will be necessary to elucidate this relationship.

Future research will also benefit from exploring the potential reinforcing effect of different modes of self-regulatory assessment. Researchers may wish to empirically test whether the predictive relationship between self-regulation and behavior concerns is explained through the reinforcement available during assessment. Finally, the importance of self-regulation across school readiness constructs (e.g., academic, behavioral, social) may instead warrant investigation into the utility of assessment of self-regulation itself, rather than as a proxy for behavioral concerns.

This study attempted to conceptualize self-regulation as an underlying mechanism of disruptive behavior, analogous to conducting screening assessment for reading fluency to predict risk for reading difficulties. However, whereas early phonemic skills serve as the "building blocks" of later reading (e.g., Goffreda et al., 2009), it does not appear self-regulation is as clearly related to externalizing problem behaviors as this study proposed. Indeed, self-regulation's importance permeates

across behavioral, social, and academic functioning, and this study provides evidence it may not be a panacea for assessing behavioral outcomes.

ENDNOTES

[1]The formula was obtained by first converting the standard scores (i.e., $\bar{X}$ = 100, SD = 15) into *z* scores ($\bar{X}$ = 0, SD = 1) using the formula $SS = (z * 15) + 100$ and solving for *z*. *z*-scores may be converted into *T* scores ($\bar{X}$ = 50, SD = 10) using the formula *T* = (*z* * 10) + 50. Thus, a conversion from Standard Score ($\bar{X}$ = 100, SD = 15) to T score may be achieved by solving $\frac{(SS-100)}{15} = \frac{(T-50)}{10}$ for *T*, which produces in $T = \frac{2}{3}SS - 16$.

[2]Chen and colleagues' (2009) paper only compared relative effect sizes of odds ratios greater than one for their guidelines. In order to determine the effect size of odds ratios smaller than one, those values were inversed (e.g., OR = 0.97 was converted to $\frac{1}{0.97} = 1.03$).

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232.

Achenbach, T. M. (1991a). *Integrative guide for the 1991 CBCL/4–18, YSR, and TRF profiles.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Child Behavior Checklist/4–18 and 1991 profile.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2000). *Achenbach System of Empirically Based Assessment.* Burlington, VT: ASEBA Research Center for Children, Youth, and Families.

Bell, M. A., & Deater-Deckard, K. (2007). Biological systems and the development of self-regulation: Integrating behavior, genetics, and psychophysiology. *Journal of Developmental and Behavioral Pediatrics, 28*(5), 409-420. doi: http://dx.doi.org/10.1097/DBP.0b013e3181131fc7

Bell, B. A., Ene, M., Smiley, W., & Shonenberger, J. A. (2013). A multilevel primer using SAS® PROC MIXED. In *SAS Global Forum 2013 Proceedings, Paper 433–2013.* Retrieved from http://support.sas.com/resources/papers/proceedings13/433-2013.pdf

Berdan, L. E., Keane, S. P. and Calkins, S. D. (2008) Temperament and externalizing behavior: social preference and perceived acceptance as protective factors. *Developmental Psychology, 44*(4), 957-968. doi: http://dx.doi.org/10.1037/0012-1649.44.4.957

Blair, C., & Peters, R. (2003). Physiological and neurocognitive correlates of adaptive behavior in preschool among children in Head Start. *Developmental Neuropsychology, 24*(1), 479–497. doi:10.1207/S15326942DN2401_04

Blair, C., & Peters Razza, R. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in Kindergarten. *Child Development, 78*(2), 647-663.

Breitenstein, S. M., Hill, C., & Gross, D. (2009). Understanding disruptive behavior problems in preschool children. *Journal of Pediatric Nursing, 24*(1), 3-12.

Bronson, M. B. (2000). *Self-regulation in early childhood: Nature and nurture.* New York, NY: Guilford Press.

Bulmer, M. G. (1979). *Principles of statistics.* New York, NY: Dover.

Burke, M. D., Rispoli, M., Clemens, N. H., Lee, Y., Sanchez, L., & Hatton, H. (2016). Integratinguniversal behavioral screening within program-wide positve behavioral interventions and supports. *Journal of Positive Behavior Interventions, 18*(1), 5-16. doi: 10.1177/1098300715580993

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. doi: https://doi.org/10.1111/1467-8624.00333

Caughy, M. O., Mills, B., Owen, M. T., & Hurst, J. R. (2013). Emergent self-regulation skills among very young ethnic minority children: A confirmatory

factor model. *Journal of Experimental Child Psychology, 116*(4), 839–855. doi: https://doi.org/10.1016/j.jecp.2013.07.017

Chen, H., Cohen, P., & Chen, P. (2009). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics – Simulation and Computation, 39*(4), 860–864. doi: https://doi.org/10.1080/03610911003650383

Conners, K. C. (2009). *Conners Early Childhood.* North Tonawanda, NY: Multi-Health Systems, Inc.

Denham, S. A., Bassett, H. H., Sirotkin, Y. S., Brown, C., & Morris, C. S. (2015). "No-o-o-o peeking": Preschoolers' executive control, social competence, and classroom adjustment. *Journal of Research in Childhood Education, 29*(2), 212-225. doi: 10.1080/02568543.2015.1008659

DiStefano, C. A., & Kamphaus, R. W. (2007). Development and validation of a behavioral screen for preschool-age children. *Journal of Emotional and Behavioral Disorders, 15*(2), 93–102. doi: 10.1177/10634266070150020401

Doll, B., & Cummings, J. (2008). Best practices in population-based school mental health services. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1333–1347). Bethesda, MD: National Association of School Psychologists

Eisenberg, N., Spinrad, T. L., Fabes, R., Reiser, M., Cumberland, A., Shepard, S. A., . . . Thompson, M. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development, 75*(1), 25–46.

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2011). Emotion-related self-regulation and its relation to children's maladjustment. *Annual Review of Clinical Psychology, 6*(1), 495–525. doi:10.1146/annurev.clinpsy.121208.131208

Eisenberg, N., Valiente, C., & Eggum, N. D. (2010). Self-regulation and school readiness. *Early Education and Development, 21*(5), 681-698. doi: 10.1080/10409289.2010.497451

Espy, K. A., Sheffield, T. D., Wiebe, S. A., Clark, C. A. C., & Moehr, M . J. (2011). Executive control and dimensions of problem behaviors in preschool children. *Journal of Child Psychology and Psychiatry, 52*(1), 33–46. doi: 10.1111/j.1469-7610.2010.02265.x

Ferrier D. E., Bassett, H. H., & Denham, S. A. (2014). Relations between executive function and emotionality in preschoolers: Exploring a transitive cognition–emotion linkage. *Frontiers in Psychology, 5*, 1-12. doi:10.3389/fpsyg.2014.00487

George, D. and Mallery, P. (2010). *SPSS for Windows Step by Step A Simple Guide and Reference 17.0 Update* (10th Ed.). Boston, MA: Pearson.

Goffreda, C. T., Diperna, J. C., & Pedersen, J. A. (2009). Preventive screening for early readers: Predictive validity of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). *Psychology in the Schools, 46*(6), 539-553. doi: https://doi.org/10.1002/pits.20396

Goldsmith, H. H., Reilly, J., Lemery, K. S., Longley, S., & Prescott, A. (1993). *Preschool Laboratory Temperament Assessment Battery (PS Lab-TAB;*

Version 1.0) [Technical report]. Department of Psychology, University of Wisconsin–Madison.

Graziano, P. A., Slavec, J., Ros, R., Garb, L., Hart, K., & Garcia, A. (2015). Self-Regulation Assessment Among Preschoolers With Externalizing Behavior Problems. *Psychological Assessment, 27*(4), 1337-1348. doi: http://dx.doi.org/10.1037/pas0000113

Gresham, F. M. (2007). Response to intervention and emotional and behavioral disorders: Best practices in assessment for intervention. *Assessment for Effective Intervention, 32*(4), 214-222. doi: 10.1177/15345084070320040301

Gresham, F. M., & Elliot, S. N. (2008). *Social Skills Improvement System Rating Scales.* San Antonio, TX: Pearson.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences, 13*(2), 65–73.

Hofman, W., Schmeichel, B. J., & Baddeley, A. D. (2012. Executive functions and self-regulation. *Trends in cognitive sciences, 16*(3), 174-180. doi:10.1016/j.tics.2012.01.006

Hosmer, D. W., Lemeshow, S. & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Hughes, C., & Ensor, R. J. (2008). Does executive function matter for preschoolers' problem behaviors? *Journal of Abnormal Child Psychology, 36*(1), 1–14. doi: 10.1007/s10802-007-9107-6

Kalberg, J. R., Lane, K. L., Driscoll, S., & Wehby, J. (2011). Systematic screening for emotional and behavioral disorders at the high school level: A formidable and

necessary task. *Remedial and Special Education, 32*(6), 506–520. doi:10.1177/0741932510362508

Kamphaus, R. W., Reynolds, C. R., & Dever, B. V. (2014). Behavioral and mental health screening. In R. J. Kettler, T. A. Glover, C. A. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational Settings: Evidence-based decision making for schools.* Washington, DC: American Psychological Association. doi: http://dx.doi.org/10.1037/14216-010

Kilgus, S. P., & Eklund, K. R. (2016). Consideration of base rates within universal screening for behavioral and emotional risk: A novel procedural framework. *School Psychology Forum: Research in Practice, 10*(1), 120-130.

Kilgus, S. P., Reinke, W. M., & Jimerson, S. R. (2015). Understanding mental health intervention and assessment within a multi-tiered framework: Contemporary science, practice, and policy. *School Psychology Quarterly, 30*(2), 159-165. doi: http://dx.doi.org/10.1037/spq0000118

Kochanska, G., Coy, K. C., & Murray, K. T. (2001). The development of self-regulation in the first four years of life. *Child Development, 72*(4), 1091–1111. doi: https://www.jstor.org/stable/1132431

Konold, T. R., Walthall, J. C., & Pianta, R. C. (2004). The behavior of child behavior ratings: Measurement structure of the Child Behavior Checklist across time, informants, and child gender. *Behavioral Disorders, 29*(4), 372–383. doi: 10.1177/019874290402900405

Kristensen, S., Henriksen, T. B., & Bilenberg, N. (2010). Assessment and analysis of parent- and caregiver-reported problems in a population-based sample of

Danish preschool children. *Nordic Journal of Psychiatry, 64*(3), 203–209. doi: 10.3109/08039480903456595

Lane, K. L., Little, M. A., Casey, A. M., Lambert, W., Wehby, J., & Weisenbach, J. L. (2009). A comparison of systematic screening tools for emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 17*(1), 93-105. doi: https://doi.org/10.1177/1063426608326203

Lane, K. L., Menzies, H. M., Bruhn, A. L., & Crnobori, M. (2011). *Managing challenging behaviors in schools: Research-based strategies that work.* New York, NY: Guilford Press.

Lane, K. L., Menzies, H. M., Oakes, W. P., & Kalberg, J. R. (2012). *Systematic screenings of behavior to support instruction: From preschool to high school.* New York, NY: Guilford Press.

Lane, K. L., Oakes, W. P., Harris, P. J., Menzies, H. M., Cox, M., & Lambert, W. (2012). Initial evidence for the reliability and validity of the student risk screening scale for internalizing and externalizing behaviors at the elementary level. *Behavioral Disorders*, *37*(2), 99–122. Retrieved from http://www.jstor.org/stable/23890734

Lin, H., Lawrence, F. R., & Gorrell, J. (2003). Kindergarten teachers' views of children's readiness for school. *Early Childhood Research Quarterly, 18*(2), 225-237. doi: https://doi.org/10.1016/S0885-2006(03)00028-0

Loeber, R., Burke, J. D., Lahey, B. B., Winters, A., & Zera, M. (2000). Oppositional defiant and conduct disorder: A review of the past 10 years, part I. *Journal of*

the American Academy of Child and Adolescent Psychiatry, 39(12), 1468−1484. doi: 10.1097/00004583-200012000-00007

Loeber, R., & Farrington, D. P. (2000). Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. *Development and Psychopathology, 12*(4), 737−762.

Lonigan, C. J., Phillips, B. M., & Hooe, E. S. (2003). Relations of positive and negative affectivity to anxiety and depression in children: evidence from a latent variable longitudinal study. *Journal of consulting and clinical psychology*, *71*(3), 465-481. http://dx.doi.org/10.1037/0022-006X.71.3.465

McClelland, M. M., & Cameron, C. E. (2012). Self-Regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development .Perspectives*, *6*(2), 136–142. doi:10.1111/j.1750-8606.2011.00191.x

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*(4), 947–959. doi: 10.1037/0012-1649.43.4.947

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., & Pratt, M. E. (2014). Predictors of early growth in academic achievement: The Head-Toes-Knees-Shoulders task. *Frontiers in Psychology, 5*, 1–14. doi:10.3389/fpsyg.2014.00599

McConaghy, S. H., & Ritter, D. R. (1995). Multidimensional assessment of emotional or behavioral disorders. In A. Thomas & J. Grimes (Eds.), *Best practices in*

school psychology–III (pp. 865- 878). Washington, DC: National Association of School Psychologists.

McMahon, R. J. (1994). Diagnosis, assessment, and treatment of externalizing problems in children: The role of longitudinal data. *Journal of Consulting and Clinical Psychology*, *62*(5), 901–917. doi:http://dx.doi.org/10.1037/0022-006X.62.5.901

Menard, S. (1995). *Applied logistic regression analysis (Sage University Paper Series on Quantitative Applications in the Social Sciences, 07–106).* Thousand Oaks, CA: Sage.

Mental Health America. (2017). *Recognizing mental health problems in children.* Retrieved from http://www.mentalhealthamerica.net/recognizing-mental-health-problems-children

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Chi, L., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication—Adolescent supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(10), 980–989. doi: 10.1016/j.jaac.2010.05.017

Merrell, K. W. (2000). Informant report: Rating scale measures. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessment of child and adolescent behaviors* (pp. 203–234). New York: Guilford.

Metcalfe, L. A., Harvey, E. A., & Laws, H. B. (2013). The longitudinal relation between academic/cognitive skills and externalizing behavior problems in

preschool children. *Journal of Educational Psychology*, *105*(3), 881–894. doi:10.1037/a0032624

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*(4907), 933–938. doi: 10.1126/science.2658056

Miyake, A. U., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. doi:10.1006/cogp.1999.0734

Montes, G., Lotyczewski, B. S., Halterman, J. S., & Hightower, A. D. (2012). School readiness among children with behavior problems at entrance into kindergarten: Results from a US national study. *European Journal of Pediatrics*, *171*(3), 541–548. doi:10.1007/s00431-011-1605-4

Montroy, J. J., Bowles, R. P., Skibbe, L. E., & Foster, T. D. (2014). Social skills and problem behaviors as mediators of the relationship between behavioral self-regulation and academic achievement. *Early Childhood Research Quarterly, 29*(3), 298–309. http://dx.doi.org/10.1016/j.ecresq.2014.03.002

NICHD Early Child Care Research Network. (2003). Do children's attention processes mediate the link between family predictors and school readiness? *Developmental Psychology, 39*(3), 581–593. doi: http://dx.doi.org/10.1037/0012-1649.39.3.581

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental*

Neuropsychology, *1*(4), 349-370. doi: http://dx.doi.org/10.1080/87565648509540320

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*(3), 605–619. doi: 10.1037/a0015365.

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment, 87*(1), 102-112. doi: 10.1207/s15327752jpa8701_09

Reynolds, C. R., & Kamphaus, R. W. (2004). *BASC-2: Behavior Assessment System for Children* (2nd ed.). Circle Pines, MN: American Guidance Service.

Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. Early Childhood Research Quarterly, *15*(2), 147–166. doi:10.1016/S0885-2006(00)00049-1

Rothbart, M. K. (2011). *Becoming who we are: Temperament and personality in development*. New York, NY: Guilford Press.

Rothbart, M. K., & Bates, J. E. (2006). Temperament. In W. Damon, R. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (6th ed., Vol. 3, pp. 99–106). New York, NY: Wiley.

Rudasill, K. M., Prokasky, A., Tu, X., Frohn, S. Sirota, K., & Molfese, V. J. (2014). Parent vs. teacher ratings of children's shyness as predictors of language and attention skills. *Learning and Individual Differences, 34,* 57-62. doi: https://doi.org/10.1016/j.lindif.2014.05.008

SAS Institute. (2018). *Statistical Analysis Software Version 9.4, University Edition.* Cary, NC: SAS Institute Inc.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods, 12*(2), 347–367. https://doi.org/10.1177/1094428107308906

Shaul, S., & Schwartz, M. (2014). The role of the executive functions in school readiness among preschool-age children. *Reading and Writing, 27*(4), 749–768. doi:10.1007/s11145-013-9470-3

Snyder, H. (2001). Epidemiology of official offending. In R. Loeber (Ed.), *Child delinquents: Development, intervention, and service needs* (pp. 25–46). Thousand Oaks, CA: Sage.

Sugai, G., & Horner, R. H. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*(2), 245–259.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better decisions through science. *Scientific American, 283*(4), 82–87. DOI: 10.1038/scientificamerican1000-82

Tape, T. G. (2017). *Plotting and interpreting an ROC curve.* Retrieved from http://gim.unmc.edu/dxtests/ROC2.htm

U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs. (2015). *37th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2015.* Washington, DC: Author.

White, B. A., Jarrett, M. A., & Ollendick, T. H. (2013). Self-regulation deficits explain the link between reactive aggression and internalizing and externalizing behavior problems in children. *Journal of Psychopathology and Behavioral Assessment, 35*(1), 1-9. doi: https://doi.org/10.1007/s10862-012-9310-9

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology, 44*(2), 575–587. doi:10.1037/0012-1649.44.2.575

Wiebe, S. A., Sheffield, T. D., & Espy, K. A. (2012). Separating the fish from the sharks: A longitudinal study of response inhibition. *Child Development, 83*(4), 1245-1261. doi: 10.1111/j.1467-8624.2012.01765.x

Wilcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention deficit/hyperactivity disorder: a metaanalytic review. *Biological Psychiatry, 57*(11), 1336–1346. doi: https://doi.org/10.1016/j.biopsych.2005.02.006

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., & Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Developmental Neuropsychology, 36*(2), 162-180. doi: 10.1080/87565641.2010.549980

APPENDIX A: *HEAD TOES KNEES SHOULDERS SCRIPT*

HTKS TASK SCRIPT

Administer the task while seated; the child should stand, about 3 feet from you, throughout the entire task. The person symbol indicates to demonstrate the correct body motions.

If the child produces the correct response immediately, score the item "2". If they self-correct right away, without prompting, score the item "1". If they do not touch the correct part of their body at all, score the item "0".

Copy Practice:

Now we're going to play a game. The game has two parts. First, I want you to copy what I do.
Touch your head.
Wait for the child to put BOTH his/her hands on head.

Good! Now touch your toes.
Wait for the child to put his/her hands on toes.

Good!
Repeat the two commands with motions again, or until the child imitates you correctly. ***(keep having child copy)***

Touch your head.

Touch your toes.

HTKS RECORD FORM

If the child produces the correct response immediately, score the item "2". If they self-correct ("see bottom of page 2) right away, without prompting, score the item "1". If they do not touch the correct part of their body at all, score the item "0".

Part 1 TRAINING: (circle child's response)

	Retraining
A1. What do you do if I say "touch your head"? 0 (head) 1 2 (toes)	_____
A2. What do you do if I say "touch your toes"? 0 (toes) 1 2 (head)	_____

PART I PRACTICE: (circle child's response)

	Incorrect	Self-Correct*	Correct	
B1. Touch your head	0 (head)	1	2 (toes)	Retraining
B2. Touch your toes	0 (toes)	1	2 (head)	
B3. Touch your head	0 (head)	1	2 (toes)	_____
B4. Touch your toes	0 (toes)	1	2 (head)	

****Retraining occurs only 3 times****

PART II TRAINING:

Administer Part II if child responds correctly to 5 or more items on Part I of the task, or if child is in kindergarten or beyond.

Circle child's response:

C1. What do you do if I say "touch your knees?" **0 (knees)** 1 **2 (shoulders)**	**Retraining** _____

PART II PRACTICE:

	Incorrect	Self-Correct*	Correct	
D1. Touch your knees	**0 (knees)**	1	**2 (shoulders)**	**Retraining**
D2. Touch your shoulders	**0 (shoulders)**	1	**2 (knees)**	
D3. Touch your knees	**0 (knees)**	1	**2 (shoulders)**	_____
D4. Touch your shoulders	**0 (shoulders)**	1	**2 (knees)**	

PART I TESTING: (circle child's response)

		Incorrect	Self-Correct*	Correct
21.	Touch your head	0 (head)	1	2 (toes)
22.	Touch your toes	0 (toes)	1	2 (head)
23.	Touch your toes	0 (toes)	1	2 (head)
24.	Touch your head	0 (head)	1	2 (toes)
25.	Touch your toes	0 (toes)	1	2 (head)
26.	Touch your head	0 (head)	1	2 (toes)
27.	Touch your head	0 (head)	1	2 (toes)
28.	Touch your toes	0 (toes)	1	2 (head)
29.	Touch your head	0 (head)	1	2 (toes)
30.	Touch your toes	0 (toes)	1	2 (head)

Total Points:_____

Number of 1 responses:_____

NOTE
***Definition of self-correction: Mark "self-correct" on both the training and testing portion if the child makes *any discernible* motion toward the *incorrect* answer, but then changes his/her mind and makes the correct response. Pausing to think, not moving, and then responding correctly does *not* count as a self-correction.**

PART II PRACTICE:

D1. Touch your knees D2. Touch your shoulders D3. Touch your knees D4. Touch your shoulders	You may use any of the remaining retraining (up to 3 total on both rules and practice) on the practice: Remember, when I say to touch your knees **(shoulders)**, you touch your shoulders **(knees)**, so you are doing something different from what I say. Let's try again.

- **If the child gets two or fewer correct**, say:

Remember, I want you to keep doing the opposite from what I say, but this time, touch your knees and shoulders.

Proceed to Part II test section. Do not explain any parts of the task again.

PART II TESTING: (circle child's response)

		Incorrect	Self-Correct	Correct
31.	Touch your head	0	1	2 (toes)
32.	Touch your toes	0	1	2 (head)
33.	Touch your knees	0	1	2 (shoulders)
34.	Touch your toes	0	1	2 (head)
35.	Touch your shoulders	0	1	2 (knees)
36.	Touch your head	0	1	2 (toes)
37.	Touch your knees	0	1	2 (shoulders)
38.	Touch your knees	0	1	2 (shoulders)
39.	Touch your shoulders	0	1	2 (knees)
40.	Touch your toes	0	1	2 (head)

Total Points:_____

Number of 1 responses:_____

PART II TESTING:

Now that you know all the parts, we're going to put them together. You're going to keep doing the opposite from what I say to do, but you won't know what I'm going to say.

There are four things I could say.

If I say to touch your head, you touch your toes.
If I say to touch your toes, you touch your head.
If I say to touch your knees, you touch your shoulders.
If I say to touch your shoulders, you touch your knees.

Are you ready? Let's try it.

11.	Touch your head
12.	Touch your toes
13.	Touch your knees
14.	Touch your toes
15.	Touch your shoulders
16.	Touch your head
17.	Touch your knees
18.	Touch your knees
19.	Touch your shoulders
20.	Touch your toes

After the child completes the task, say:
Thank you for playing this game with me today!

HTKS SCORING

Each item is coded as follows (Ponitz et al., 2008):

0 = Incorrect response
1 = Any motion to incorrect response, but self-corrected to end with correct response
2 = Correct response

Final Score:

The task has begins with 6 practice items and between the first and second set of items there are 5 more practice trials. The final score is the sum of the first six practice items and the 20 test items. (Range: 0-52)

APPENDIX B: *ACHENBACH CAREGIVER-TEACHER REPORT (C-TRF)*
EXTERNALIZING PROBLEM SCALE

Scale
For each item that describes the child *now or within the past 2 months,* please circle:

0	1	2
Not True (as far as you know)	Somewhat or Sometimes True	Very True or Often True

Items	**Attention**
5	Can't concentrate, can't pay attention for long
6	Can't sit still, restless, or hyperactive
24	Difficulty following directions
48	Fails to carry out assigned tasks
51	Fidgets
56	Poorly coordinated or clumsy
59	Quickly shifts from one activity to another
64	Inattentive, easily distracted
95	Wanders away
	Aggression
8	Can't stand waiting; wants everything now
14	Cruel to animals
15	Defiant
16	Demands must be met immediately
17	Destroys his/her own things
18	Destroys property belonging to others
20	Disobedient
22	Cruelty, bullying, or meanness to others
27	Doesn't seem to feel guilty after misbehaving
28	Disturbs other children
29	Easily frustrated
35	Gets in many fights
40	Hits others
42	Hurts animals or people without meaning to
44	Angry moods
53	Physically attacks people
58	Punishment doesn't change his/her behavior
66	Screams a lot
69	Selfish or won't share
74	Not liked by other children
81	Stubborn, sullen, or irritable
84	Teases a lot
85	Temper tantrums or hot temper
88	Uncooperative
96	Wants a lot of attention-

APPENDIX C: CONNERS EARLY CHILDHOOD BEHAVIOR FORM-TEACHER REPORT (EC BEH-T) DEFIANT/AGGRESSIVE BEHAVIORS TOTAL SCALE

Scale
In the past month, this was…

0	1	2	3
Not true at all (Never, Seldom)	Just a little true (Occasionally)	Pretty much true (Often, Quite a bit)	Very much true (Very often, Very frequently)

Items	
9	Steals
20	Temper outbursts; explosive, unpredictable behavior
40	Lies to get things or to manipulate people
46	Swears or uses bad language
53	Sulks
57	Tries to hurt other people's feelings
72	Is cold-hearted and cruel
73	Is rude
75	Mood changes quickly and drastically
76	Gets into fights
86	Threatens people
98	Picks on other children
101	Is defiant
102	Destroys things on purpose
104	Is cruel to animals
105	Argues with adults
108	Is manipulative
111	Is bossy

APPENDIX D: *SOCIAL SKILLS IMPROVEMENT SYSTEM RATING SCALES TEACHER FORM* (*SSIS*) EXTERNALIZING SUBSCALE

Scale

Please read each item and think about this student's behavior during the past two months. Then, decide **how often** this student displays this behavior:

N	S	O	A
Never	Seldom	Often	Almost Always

Items

47	Acts without thinking
49	Bullies others
51	Has difficulty waiting for turn
53	Fidgets or moves around too much
55	Forces others to act against their will
57	Has temper tantrums
61	Is aggressive toward people or objects
63	Cheats in games or activities
67	Fights with others
69	Disobeys rules or requests
73	Talks back to adults
75	Lies or does not tell the truth

APPENDIX E: PARENT DEMOGRAPHIC QUESTIONNAIRE

Parent Input and Information

Directions: We would like you to complete the following items about you and your child. When filling out this information, please use the following as a guide for filling in the circles correctly:

All information you provide will be kept confidential. There are no right or wrong answers to questions. The information you provide will help us better understand you and your child. Any information provided on this form will be kept private.

Your name: _____

Today's date: _____

1. What is your relationship to the child? (**choose one**)
   - o Mother
   - o Father
   - o Grandmother
   - o Grandfather
   - o Stepmother
   - o Stepfather
   - o Foster mother
   - o Foster father
   - o Other, Please Specify: _____

2. What is your gender?
   - o Female
   - o Male

3. What is your date of birth? _____/_____/_____ (month/ day/ year)

4. What is the highest level of education you have completed? **(choose one)**
   - o No formal schooling
   - o Less than 9th grade
   - o 9th grade to 12th grade, no diploma or GED
   - o High school diploma
   - o GED
   - o Some college, but not a degree
   - o Vocational/technical training or certificate
   - o Associate's or two year college degree
   - o Four year college degree (BA, BS)
   - o Some graduate college coursework
   - o Graduate (MS, MA, PhD) or Professional degree (MD, JD, DDS)

5. What is your current marital or partner status? **(choose one)**
   - o Married
   - o In a registered domestic partnership or civil union
   - o Living with a partner
   - o Separated
   - o Divorced
   - o Widowed
   - o Never Married

6. Is there another primary caregiver in the home?
   - o Yes*
   - o No

*6a. *If yes, what is his/her relationship to the child?*
   - o Mother
   - o Father
   - o Grandmother
   - o Grandfather
   - o Stepmother
   - o Stepfather
   - o Foster mother
   - o Foster father
   - o Other, Please Specify: _____

*6b. *What is the highest level of education that caregiver has completed?* (**choose one**)
   - o No formal schooling
   - o Less than 9th grade
   - o 9th grade to 12th grade, no diploma or GED
   - o High school diploma
   - o GED
   - o Some college, but not a degree
   - o Vocational/technical training or certificate
   - o Associate's or two year college degree
   - o Four year college degree (BA, BS)
   - o Some graduate college coursework
   - o Graduate (MS, MA, PhD) or Professional degree (MD, JD, DDS)

7. What was the total combined income of all members of your household in the last calendar year? Please include income from jobs, businesses, child support, welfare, social security/ disability payment, alimony, unemployment, pensions, dividends, and any other money.

- o $0
- o $1 - $5,000
- o $5,001 - $10,000
- o $10,001 - $15,000
- o $15,001 - $20,000
- o $20,001 - $25,000
- o $25,001 - $30,000
- o $30,001 - $35,000
- o $35,001 - $40,000
- o $40,001 - $45,000
- o $45,001 - $50,000
- o $50,001 - $55,000
- o $55,001 - $60,000
- o $60,001 - $65,000
- o $65,001 - $70,000
- o $70,001 - $75,000
- o $75,001 - $100,000
- o $100,001 - $200,000
- o $200,001 or more
- o Prefer not to answer
- o I don't know

Please answer the following questions about your child in this study.

1. What is your child's gender?

- o Female
- o Male

2. What is your child's birth date? _____/_____/_____ (month/ day/ year)

3. Is your child Hispanic or Latino? **(choose one)**

- o No, not Hispanic or Latino
- o Yes, Mexican, Mexican American, Chicano
- o Yes, Puerto Rican
- o Yes, Cuban
- o Yes, another Hispanic, Latino, or Spanish origin -- **Please print origin**, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)

_____

4. What is your child's race? (Regardless of how you answered previous question, **choose one or more**)

- American Indian or Alaska Native
- Asian or Asian American
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White
- Some other race – Please specify: _____
- Prefer not to answer.

5. Have you ever had a concern about delays or differences in your child's development?

- Yes*
- No

*5a. *If yes, what was the concern?* _____

6. Has a health care provider, childcare provider or other professional stated concerns about delays or differences in your child's development?

- Yes*
- No

*6a. *If Yes, what was the concern?* _____\

7. Does your child currently have an Individualized Education Plan (IEP)?

- Yes
- No
- I don't know

8. How many different care arrangements, other than home, does your child spend at least 10 hours in per week? _____ (Number of different arrangements, including preschool and childcare centers)

9. In the **prior school year** (September 2016 – June 2017), where did your child spend his/her time during daytime hours? **(Select all that apply)**

- Head Start program (Head Start is a federally sponsored pre-kindergarten program primarily for children from low income families)
  Specify name:_____
- Preschool in a public school
  Specify name:_____
- An early education center, child care center, parochial child care center, or nursery school other than Head Start
  Specify name:_____
- An in-home child care program or family child care program
  Specify name:_____
- An "extended-day" program, that is, before- or after-school care at the child's regular school

- Care by a parent
- Care by a member of your family or household
- Care by someone other than a member of your family or household
- Other, please specify: _____

10 . What is the name of the place that your child spent the most time in the prior school year?

Specify name: _____

11. In the **summer of 2017** (July 2016 - August 2016), where did your child spend his/her time during daytime hours? **(Select all that apply)**

- Head Start program (Head Start is a federally sponsored pre-kindergarten program primarily for children from low income families)
  Specify name:_____
- Preschool in a public school
  Specify name:_____
- An early education center, child care center, parochial child care center, or nursery school other than Head Start
  Specify name:_____
- An in-home child care program or family child care program
  Specify name:_____
- An "extended-day" program, that is, before- or after-school care at the child's regular school
- Care by a parent
- Care by a member of your family or household
- Care by someone other than a member of your family or household
- Other, please specify: _____
-

17. What is the name of the place that your child spent the most time during the summer of 2017?
    Specify name: _____

18. In the **fall of 2017** (September 2017 – December 2017), where did your child spend his/her time during daytime hours? **(Select all that apply)**
    - Head Start program (Head Start is a federally sponsored pre-kindergarten program primarily for children from low income families)
      Specify name:_____
    - Preschool in a public school
      Specify name:_____
    - An early education center, child care center, parochial child care center, or nursery school other than Head Start
      Specify name:_____
    - An in-home child care program or family child care program
      Specify name:_____
    - An "extended-day" program, that is, before- or after-school care at the child's regular school
    - Care by a parent
    - Care by a member of your family or household
    - Care by someone other than a member of your family or household
    - Other, please specify: _____

19. What is the name of the place that your child spent the most time during the fall of 2017?
Specify name: _____

20. Did your child attend any structured pre-kindergarten programs ***before August, 2016***?
    - o Yes*
    - o No

*20a. *If yes, how many total months did your child attend pre-kindergarten programs in each of the following settings* ***before August, 2016****?*

- o Head Start program (Head Start is a federally sponsored pre-kindergarten program primarily for children from low income families)
  _____ months
- o Preschool in a public school
  _____ months
- o An early education center, child care center, parochial child care center, or nursery school other than Head Start
  _____ months
- o An in-home child care program or family child care program
  _____ months
- o An "extended-day" program, that is, before- or after-school care at the child's regular school
  _____ months
- o Other, please specify: _____

  _____ months

APPENDIX F: SCRIPT FOR OBTAINING CHILD ASSENT

Hello, my name is _____ and I work with the University of Nebraska. Your parent and teacher gave me permission to play a quick game with you for a research project I am working on. If you finish the game, you will be able to choose a prize for your time. You do not have to play the game and can go back to your class if you wish. Do you have any questions?

The game will take about five minutes to play, would you like to begin?

[IF AGREE, BEGIN HTKS SCRIPT]

APPENDIX G GROUP MEAN COMPARISONS

Group Mean Differences for Covariates and Head Toes Knees Shoulders

ANOVA for Mother's Education on HTKS

	Sum of Squares	df	Mean Square	F
Model	374.13	4	93.53	0.37
Error	1555.70	19	81.88	
Total	1929.83	23		

*p < 0.05

ANOVA for Parent Income on HTKS

	Sum of Squares	df	Mean Square	F
Model	534.90	5	106.98	1.38
Error	1394.93	18	77.50	
Total	1929.83	23		

*p < 0.05

ANOVA for Previous Preschool Experience on HTKS

	Sum of Squares	df	Mean Square	F
Model	130.66	4	32.67	0.34
Error	1799.17	19	94.69	
Total	1929.83	23		

*p < 0.05

ANOVA for Ethnic Minority Status on HTKS

	Sum of Squares	df	Mean Square	F
Model	4.09	1	4.09	0.04
Error	1921.91	21	91.52	
Total	1926.00	22		

*p < 0.05

ANOVA for Gender on HTKS

	Sum of Squares	df	Mean Square	F
Model	46.21	1	46.21	0.52
Error	1879.79	21	86.51	
Total	1926.00	22		

*p < 0.05

Group Mean Differences for Covariates and the CBCL – *C-TRF*

ANOVA for Mother's Education on C-TRF Externalizing Problems

	Sum of Squares	df	Mean Square	F
Model	604.42	4	151.1	1.53
Error	1581.58	16	98.85	
Total	2186.00	20		

*p < 0.05

ANOVA for Parent Income on C-TRF Externalizing Problems

	Sum of Squares	df	Mean Square	F
Model	711.53	5	142.31	1.45
Error	1474.47	15	98.30	
Total	2186.00	20		

*p < 0.05

ANOVA for Previous Preschool Experience on C-TRF Externalizing Problems

	Sum of Squares	df	Mean Square	F
Model	957.57	4	239.39	3.12*
Error	1228.43	16	76.78	
Total	2186.00	20		

*p < 0.05

ANOVA for Ethnic Minority Status on C-TRF Externalizing Problems

	Sum of Squares	df	Mean Square	F
Model	6.56	1	6.56	0.06
Error	2179.44	19	114.71	
Total	2186.00	20		

*p < 0.05

ANOVA for Gender on C-TRF Externalizing Problems

	Sum of Squares	df	Mean Square	F
Model	5.83	1	5.83	0.05
Error	2180.17	19	114.76	
Total	2186.00	20		

*p < 0.05

Group Mean Differences for Covariates and Conners *EC BEH-T*

ANOVA for Mother's Education on EC BEH-T Defiant/Aggressive Behaviors

	Sum of Squares	df	Mean Square	F
Model	771.53	4	192.88	0.45
Error	3199.71	16	199.98	
Total	3971.24	20		

*p < 0.05

ANOVA for Parent Income on EC BEH-T Defiant/Aggressive Behaviors

	Sum of Squares	df	Mean Square	F
Model	1804.47	5	360.91	0.08
Error	2166.67	15	144.44	
Total	3971.24	20		

*p < 0.05

ANOVA for Previous Preschool Experience on EC BEH-T Defiant/Aggressive Behaviors

	Sum of Squares	df	Mean Square	F
Model	1042.31	4	260.58	1.42
Error	2928.93	16	183.06	
Total	3971.24	20		

*p < 0.05

ANOVA for Ethnic Minority Status on EC BEH-T Defiant/Aggressive Behaviors

	Sum of Squares	df	Mean Square	F
Model	19.50	1	19.50	0.09
Error	3951.74	19	208.00	
Total	3971.24	20		

*p < 0.05

ANOVA for Gender on EC BEH-T Defiant/Aggressive Behaviors

	Sum of Squares	df	Mean Square	F
Model	18.30	1	18.30	0.09
Error	3953.94	19	208.05	
Total	3971.24	20		

*p < 0.05

Group Mean Differences for Covariates and *SSIS*

ANOVA for Mother's Education on SSIS Problem Behaviors

	Sum of Squares	df	Mean Square	F
Model	120.66	4	30.17	1.17
Error	411.31	16	25.71	
Total	531.98	20		

*p < 0.05

ANOVA for Parent Income on SSIS Problem Behaviors

	Sum of Squares	df	Mean Square	F
Model	200.37	5	40.07	1.81
Error	331.61	15	22.11	
Total	531.98	20		

*p < 0.05

ANOVA for Previous Preschool Experience on SSIS Problem Behaviors

	Sum of Squares	df	Mean Square	F
Model	123.08	4	30.77	1.20
Error	408.90	16	25.56	
Total	531.98	20		

*p < 0.05

ANOVA for Ethnic Minority Status on SSIS Problem Behaviors

	Sum of Squares	df	Mean Square	F
Model	7.29	1	7.29	0.26
Error	524.69	19	27.62	
Total	531.98	20		

*p < 0.05

ANOVA for Gender on SSIS Problem Behaviors

	Sum of Squares	df	Mean Square	F
Model	3.39	1	3.39	0.12
Error	528.59	19	27.82	
Total	531.98	20		

*p < 0.05